Yi-Hsuan Yang (楊奕軒)[1,2]

1 Professor
Department of Electrical Engineering
National Taiwan University

2 Joint-Appointed Research Fellow
Research Center for Information Technology Innovation (CITI),
Academia Sinica

EE2-337, Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan
https://affige.github.io/
https://scholar.google.com/citations?user=OL-XGxcAAAAJ&hl=zh-TW
affige@gmail.com

EXPERTISE

Music information research; Artificial intelligence; Machine learning; Music generation; Affective Computing

EDUCATION

- Ph.D., Communication Engineering, National Taiwan University, Taiwan 2010
- B.S., Electrical Engineering, National Taiwan University, Taiwan 2006

WORK EXPERIENCES

- **Full Professor**, Dept. Electrical Engineering, National Taiwan University **since 2023/02**
- Joint-Appointment Researcher, Research Center for IT Innovation, Academia Sinica since 2023/08
- **Chief Music Scientist**, Taiwan AI Labs **2019-2023**
- **Associate Research Professor**, Research Center for IT Innovation, Academia Sinica **2015-2023**
- Joint-Appointment Associate Professor, CSIE, National Cheng Kung University 2017-2019
- Adjunct Associate Professor, CSIE, National Tsing-Hua University 2016
- **Assistant Research Professor**, Research Center for IT Innovation, Academia Sinica **2011−2015**
- Visiting Scholar (three months)**,** Columbia University, USA 2013
- Visiting Scholar (three months), Music Technology Group, Universitat Pompeu Fabra, Spain 2011
- Second Lieutenant (one year), Communications, Electronics and Information, ROC Army 2010-2011

AWARDS & HONORS

- Outstanding Teaching Awards, National Taiwan University 2025
- NTU-Fubon Distinguished Chair Professorship Award 2023
- Multimedia Rising Stars Award, IEEE International Conference on Multimedia Expo. (ICME) 2019
- Best Associate Editor Service Award, IEEE Transactions on Multimedia 2018
- Best Conference Paper Award, IEEE Multimedia Communications Technical Committee (MMTC) 2015
- Best Paper Award, IEEE International Conference on Multimedia Expo. (ICME) 2015

- Young Scholars' Creativity Award, Foundations for the Advancement of Outstanding Scholarship 2015
- Ta-You Wu Memorial Research Award, Ministry of Science and Technology 2014
- Best Poster Award, IEEE/ACM Joint Conference on Digital Libraries 2014
- Project for Excellent Junior Research Investigators, National Science Council 2013−2016
- Career Development Award, Academia Sinica 2013−2017
- Pan Wen Yuan Research Exploration Award 2013
- First Prize, ACM Multimedia Grand Challenge 2012
- IEEE SPS Young Author Best Paper Award, IEEE Signal Processing Society 2011
- Best Ph.D. Dissertation Award, Graduate Institute of Communication Engineering, NTU 2010
- Best Ph.D. Dissertation Award, TAAI (Taiwanese Association for Artificial Intelligence) 2010
- MediaTek Fellowship 2009
- Microsoft Research Asia (MSRA) Fellowship 2008

SELECTED RECENT PUBLICATIONS

- "MuseControlLite: Multifunctional music generation with lightweight conditioners," *ICML* 2025
- "MuseMorphose: Full-song and fine-grained piano music style transfer with just one Transformer VAE," *TASLP* 2023.
- "Relative positional encoding for Transformers with linear complexity," *ICML* 2021.
- "Compound Word Transformer: Learning to compose full-song music over dynamic directed hypergraphs," *AAAI* 2021.
- "Pop Music Transformer: Beat-based modeling and generation of expressive Pop piano compositions," *ACM Multimedia* 2020.
- "Musical composition style transfer via disentangled timbre representations," *IJCAI* 2019.
- "Score-to-audio music generation with multi-band convolutional residual network," *AAAI* 2019.
- "MuseGAN: Multi-track sequential GAN for symbolic music generation and accompaniment," *AAAI* 2018.
- "Generating music medleys via playing music puzzle games," *AAAI* 2018
- "MidiNet: A convolutional GAN for symbolic-domain music generation," *ISMIR* 2017
- *Music Emotion Recognition*, CRC Taylor & Francis Books, Feb. 2011.

ACADEMIC SERVICES

- **Test-of-Time Award Committee Member** of
  - Int. Society for Music Information Retrieval Conference (ISMIR) 2025
- **Associate Editor** of
  - IEEE Transactions on Multimedia 2016/9-2019/2
  - IEEE Transactions on Affective Computing 2016/11-2019/2
- **IEEE Senior Member** since 2017
- **Program Chair** of
  - Int. Society for Music Information Retrieval Conference (ISMIR) 2014

- **Guest Editor** of
  - ACM Transactions on Intelligent Systems and Technology 2015
  - IEEE Transactions on Affective Computing 2014
- **10K Award Committee Member** of
  - IEEE International Conference on Multimedia and Expo. (ICME) 2016−2018
- **Tutorial Chair** of
  - Int. Society for Music Information Retrieval Conference (ISMIR) 2021
- **Unconference Chair** of
  - Int. Society for Music Information Retrieval Conference (ISMIR) 2017
- **External PhD thesis committee member** of
  - Hong Kong University of Science and Technology 2015
- **Senior PC Member (Meta-reviewer)** of
  - AAAI 2022, ISMIR 2021,etc
- **Organizer** of
  - ICME 2026 Grand Challenge on Academic Text-to-Music Generation 2026
  - Int. Workshop on Affect and Sentiment in Multimedia, in conjunction with ACM MM 2015
  - MediaEval Affect Task: Music in Emotion 2013−2015
  - MIREX Singing Voice Separation Task 2014−2015
  - Int. Workshop on Affective Analysis in Multimedia, in conjunction with IEEE ICME 2013
  - Taiwanese Workshop on Music Information Retrieval 2012−2014

PROJECTS

- Towards Controllable and Affordable Text-to-Music Generation MOST 2025-2028
- Lightweight and Controllable Long-form Music Generation Google Gift 2025
- Controllable and Playable Music Generation Google Gift 2024
- Human-level and Open-Source Mandarin Pop Music Melody and Singing Generation by AI MOST 2023-2025
- Open DJ Project (II): Automatic EDM Generation MOST 2020-2022
- GenMusic Project: Industrial AI-Powered Music Composition Platform MOST 2018-2020
- Open DJ Project: AI for Automatic and Personalized DJing MOST 2018-2020
- A Unified Framework for Processing and Understanding Heterogeneous Data for Intelligent Recommendation (co-PI) MOST 2017-2020
- Product Recommendation and Customer Status Prediction Cathay 2017-2018
- Online Guitar Transcription: Melody, Chord and Playing Techniques Recognition MOST 2016-2018
- Mobile Music Recommendation using Brain-Computer Interfaces MOST 2015-2018
- User-centered Intelligent Music Streaming and Recommendation Platform (III) KKBOX Inc., 2017-2019
- User-centered Intelligent Music Streaming and Recommendation Platform (II) KKBOX Inc., 2015-2017

- User-centered Intelligent Music Streaming and Recommendation Platform — KKBOX Inc., 2013-2015
- User Preference Modeling from Listening History & Artist Similarity — KKBOX Inc., 2012-2013
- Music Recommendation based on Listening Context — HTC Inc., 2012
- Dictionary-based Music Signal Analysis, Understanding, and Retrieval — Academia Sinica, 2013-2017
- Automatic Music Recommendation and Retrieval — MOST 2013-2016
- Dictionary-based Multipitch Estimation of Polyphonic Music — NSC 2012-2013
- Large-scale Music Emotion Recognition System using Social Media — NSC 2011-2012

TUTORIALS

- Hao-Wen Dong and **Yi-Hsuan Yang**, "Generating Music with GANs: An Overview and Case Studies," *Int. Society for Music Information Retrieval Conference* (*ISMIR*), 2019 (link).
- Xiao Hu and **Yi-Hsuan Yang**, "Music Affect Recognition: The State-of-the-art and Lessons Learned," *Int. Society for Music Information Retrieval Conference* (*ISMIR*), 2012.

STUDENT AWARDS

- 謝翔、王竣平、張譯心, 3rd Prize, Bachelor Thesis Award, Dept. of EE, NTU — 2025
- Yi-Jen Shih, 1st Prize, Bachelor Thesis Award, Dept. of EE, NTU — 2022
- Shih-Lun Wu, Ssu-Nien Fu's Award (1st Prize), Best Bachelor's Thesis, National Taiwan University — 2021
- Shih-Lun Wu, 1st Prize, Bachelor Thesis Award, Dept. of CSIE, NTU — 2021
- Yi-Hui Chou, 2ne Prize, Bachelor Thesis Award, Dept. of EE, NTU — 2021
- Shih-Lun Wu, 1st Prize, Bachelor Thesis Award, Dept. of CSIE, NTU — 2020
- Ching-Yu Chiu, Jury's Recommendation Award, 美律電聲論文獎 — 2021
- Yu-Hsiang Huang, Special Award, 美律電聲論文獎 — 2020
- Wen-Yi Hsiao, 2ne Prize, 美律電聲論文獎 — 2018

PUBLICATIONS

Total citations: **14,157**; *citations of most-cited paper*: **1028**; *h-index*: **59**; *i10-index*: **200**

- **Book**

[1] Y.-H. Yang and H. H. Chen, *Music Emotion Recognition*, CRC Taylor & Francis Books, Feb. 2011.

- **Proceedings (Edited)**

[2] Meinard Müller, Emilia Gómez, and Yi-Hsuan Yang, "Computational methods for melody and voice processing in music recordings," Report from Dagstuhl Seminar 19052, 2019.

[3] Hsin-Min Wang, Yi-Hsuan Yang, and Jin Ha Lee, International Society for Music Information Retrieval Conference, Proceedings, ISMIR, Taipei, Taiwan, 2014.

- **Journal Papers**

[4] Xavier Serra et al, “AI and music at a crossroads: The role of the MIR community in shaping trustworthy, open, and culturally inclusive Music-AI,” *Transactions of the International Society for Music Information Retrieval*, 2026.

[5] Yu-Hua Chen, Yuan-Chiao Cheng, Yen-Tung Yeh, Jui-Te Wu, Jyh-Shing Roger Jang, and Yi-Hsuan Yang, “Towards generalizability to tone and content variations in the transcription of amplifier rendered electric guitar audio,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 34, pp. 3000-3009, 2026.

[6] Gaël Richard, Vincent Lostanlen, Yi-Hsuan Yang, and Meinard Müller, “Model-based deep learning for music information research,” *IEEE Signal Processing Magazine*, vol. 41, no. 6, pp. 51-59, November 2024.

[7] Yi-Hui Chou, I-Chun Chen, Chin-Jui Chang, Joann Ching, and Yi-Hsuan Yang, “MidiBERT-Piano: BERT-like pre-training for symbolic piano music classification tasks," *Journal of Creative Music Systems* (JCMS), vol. 8, no. 1, 2024.

[8] Ching-Yu Chiu, Meinard Müller, Matthew E. P. Davies, Alvin Wen-Yu Su, and Yi-Hsuan Yang, "Local periodicity-based beat tracking for expressive classical piano music," *IEEE/ACM Transactions on Audio,* Speech*, and Language Processing* (TASLP), vol. 31, pp. 2824-2835, July 2023.

[9] Shih-Lun Wu and Yi-Hsuan Yang, “MuseMorphose: Full-song and fine-grained piano music style transfer with just one Transformer VAE,” *IEEE/ACM Transactions on Audio,* Speech*, and Language Processing* (TASLP), vol. 31, pp. 1953-1967, May 2023.

[10] Ching-Yu Chiu, Meinard Müller, Matthew E. P. Davies, Alvin Wen-Yu Su, and Yi-Hsuan Yang, "An analysis method for metric-level switching in beat tracking," *IEEE Signal Processing Letters* (SPL), vol. 29, 2153-2157, Oct. 2022.

[11] Yi-Jen Shih, Shih-Lun Wu, Frank Zalkow, Meinard Müller, and Yi-Hsuan Yang, “Theme Transformer: Symbolic music generation with theme-conditioned Transformer,” *IEEE Transections on Multimedia* (TMM), vol. 25, pp. 3495-3508, March 2022.

[12] Juan Sebastián Gomez-Cañón, Estefanía Cano, Tuomas Eerola, Perfecto Herrera, Xiao Hu, Yi-Hsuan Yang, and Emilia Gómez, “Music Emotion Recognition: Towards new robust standards in personalized and context-sensitive applications,” *IEEE Signal Processing Magazine*, vol. 38, no. 6, pp. 106-114, Nov. 2021.

[13] Ching-Yu Chiu, Alvin Wen-Yu Su, and Yi-Hsuan Yang, “Drum-aware ensemble architecture for improved joint musical beat and downbeat tracking,” *IEEE Signal Processing Letters* (SPL)*,* vol. 28, pp. 1100-1104, May 2021.

[14] Yin-Cheng Yeh, Wen-Yi Hsiao, Satoru Fukayama, Tetsuro Kitahara, Benjamin Genchel, Hao-Min Liu, Hao-Wen Dong, Yian Chen, Terence Leong, and Yi-Hsuan Yang, “Automatic melody harmonization with triad chords: A comparative study,” *Journal of New Music Research*, vol. 50, no. 1, pp. 37-51, 2021.

[15] E. Zangerle, C.-M. Chen, M.-F. Tsai and Y.-H. Yang, “Leveraging affective hashtags for ranking music recommendations,” *IEEE Transactions on Affective Computing* (TAC), vol. 12, no. 1, pp. 78-91, 2021.

[16] Zhe-Cheng Fan, Tak-Shing T. Chan, Yi-Hsuan Yang, and Jyh-Shing R. Jang, “Backpropagation with *N*-D vector-valued neurons using arbitrary bilinear products,” *IEEE Transactions on Neural Networks and Learning Systems* (TNNLS), vol. 31, no. 7, pp. 2638-2652, 2020.

[17] T.-W. Su, Y.-P. Chen, L. Su, and Y.-H. Yang, “TENT: Technique-embedded note tracking for real-world guitar solo recordings,” *Transactions of the International Society for Music Information Retrieval*

(TISMIR), vol. 2, no. 1, pp. 15-28, 2019.

[18] S.-Y. Chou, J.-S. R. Jang, and Y.-H. Yang, “Fast tensor factorization for large-scale context-aware recommendation from implicit feedback,” *IEEE Trans. Big Data* (TBD), vol. 6, no. 1, pp. 201-208, Mar. 2020.

[19] J.-Y. Liu, Y.-H. Yang, and S.-K. Jeng, “Weakly-supervised visual instrument-playing action detection in videos,” *IEEE Transactions on Multimedia* (TMM), vol. 21, no. 4, pp. 887-901, Apr. 2019.

[20] J. Nam, K. Choi, J. Lee, S.-Y. Chou, and Y.-H. Yang, “Deep learning for audio-based music classification and tagging,” *IEEE Signal Processing Magazine* (SPM), vol. 36, no. 1, pp. 41-51, Jan. 2019.

[21] J.-C. Lin, W.-L. Wei, T.-L. Liu, Y.-H. Yang, H.-M. Wang, H.-R. Tyan, and H.-Y. M. Liao, “Coherent deep-net fusion to classify shots in concert videos,” *IEEE Transactions on Multimedia* (TMM), vol. 20, no. 11, pp. 3123-3136, Nov. 2018.

[22] Y.-H. Chin, J.-C. Wang, J.-C. Wang and Y.-H. Yang, “Predicting the probability density function of music emotion using emotion space mapping,” *IEEE Transactions on Affective Computing* (TAC), vol. 9, no. 4, pp. 541-549, Oct.-Dec. 2018.

[23] Y.-S. Huang, S.-Y. Chou, and Y.-H. Yang, “Pop music highlighter: Marking the emotion keypoints,” *Transactions of the International Society for Music Information Retrieval* (TISMIR), vol. 1, no. 1, pp. 68-78, Sep. 2018.

[24] Y.-P. Lin, P.-K. Jao, and Y.-H. Yang, “Improving cross-day EEG-based emotion classification using robust principal component analysis,” *Frontiers in Computational Neuroscience*, Jul. 2017.

[25] A. Aljanaki, Y.-H. Yang, and M. Soleymani, “Developing a benchmark for emotional analysis of music,” *PLOS ONE*, vol. 12, no. 3, e0173392.doi:10.1371/journal.pone.0173392, Mar. 2017.

[26] X. Hu and Y.-H. Yang, “The mood of Chinese pop music: Representation and recognition,” *Journal of the Association for Information Science and Technology* (JAIST), doi:10.1002/asi.23813, Jun. 2017.

[27] Y.-A. Chen, J.-C. Wang, Y.-H. Yang, H. H. Chen, “Component tying for mixture model adaptation in personalization of music emotion recognition,” *IEEE/ACM Transactions on Audio,* Speech*, and Language Processing* (TASLP), vo. 25, no. 7, pp. 1409-1420, Jul. 2017. [cover page of the issue]

[28] X. Hu and Y.-H. Yang, “Cross-dataset and cross-cultural music mood prediction: A case on Western and Chinese pop songs,” *IEEE Transactions on Affective Computing* (TAC), vol. 8, no. 2, pp. 228-240, Apr. 2017.

[29] T.-S. Chan and Y.-H. Yang, “Informed group-sparse representation for singing voice separation,” *IEEE Signal Processing Letters* (SPL), vol. 24, no. 2, pp. 156-160, Feb. 2017.

[30] T.-S. Chan and Y.-H. Yang, “Polar *n*-complex and *n*-bicomplex singular value decomposition and principal component pursuit,” *IEEE Transactions on Signal Processing* (TSP), vol. 64, no. 24, pp. 6533-6544, Dec. 2016.

[31] M. Schedl, Y.-H. Yang, and P. Herrera, “Introduction to intelligent music systems and applications,” *ACM Transactions on Intelligent Systems and Technology* (TIST), vol. 8, no. 2, article 17, Oct. 2016.

[32] P.-K. Jao, L. Su, Y.-H. Yang and B. Wohlberg, “Monaural music source separation using convolutional sparse coding,” *IEEE/ACM Transactions on Audio,* Speech*, and Language Processing* (TASLP), vol. 24, no. 11, pp. 2158-2170, Nov. 2016.

[33] T.-S. Chan and Y.-H. Yang, “Complex and quaternionic principal component pursuit and its application

to audio separation," *IEEE Signal Processing Letters* (SPL), vol. 23, no. 2, pp. 287-291, Feb. 2016.

[34] C.-Y. Liang, L. Su and Y.-H. Yang, "Musical onset detection using constrained linear reconstruction," *IEEE Signal Processing Letters* (SPL), vol. 22, no. 11, pp. 2142-2146, Nov. 2015.

[35] L. Su and Y.-H. Yang, "Combining spectral and temporal representations for multipitch estimation of polyphonic music," *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (TASLP), vol. 23, no. 10, pp. 1600-1612, Oct. 2015.

[36] P.-K. Jao and Y.-H. Yang, "Music annotation and retrieval using unlabeled exemplars: correlation and sparse codes," *IEEE Signal Processing Letters* (SPL), vol. 22, no. 10, pp. 1771-1775, Oct. 2015.

[37] Y.-H. Yang and Y.-C. Teng, "Quantitative study of music listening behavior in a smartphone context," *ACM* Transactions *on Interactive Intelligent Systems* (TiiS), vol. 5, no. 3, article 14, Aug. 2015.

[38] M. Soleymani, Y.-H. Yang, G. Irie, and A. Hanjalic, "Challenges and perspectives for affective analysis in multimedia," *IEEE Transactions on Affective Computing* (TAC), vol. 6, no. 3, pp. 206-208, 2015.

[39] J.-C. Wang, Y.-H. Yang, H.-M. Wang, and S.-K. Jeng, "Modeling the affective content of music with a Gaussian mixture model," *IEEE Transactions on Affective Computing* (TAC), vol. 6, no. 1, pp. 56-68, Feb. 2015.

[40] L. Su, H.-M. Lin, and Y.-H. Yang, "Sparse modeling of magnitude and phase-derived spectra for playing technique classification," *IEEE Transactions on Audio, Speech, and Language Processing* (TASLP), vol. 22, no. 12, pp. 2122-2132, Dec. 2014.

[41] L. Su, C.-C. Yeh, J.-Y. Liu, J.-C. Wang, and Y.-H. Yang, "A systematic evaluation of the bag-of-frames representation for music information retrieval," *IEEE Transactions on Multimedia* (TMM), vol. 16, no. 5, pp. 1188-1200, Aug. 2014.

[42] Y.-P. Lin, Y.-H. Yang, and T.-P. Jung, "Fusion of Electroencephalogram dynamics and musical contents for estimating emotional responses in music listening," *Frontiers in Neuroscience*, vol. 8, no. 94, pp. 1-14, May 2014.

[43] Y.-H. Yang and J.-Y. Liu, "Quantitative study of music listening behavior in a social and affective context," *IEEE Transactions on Multimedia* (TMM), vol. 15, no. 6, pp. 1304-1315, Oct. 2013.

[44] K.-S. Lin, A. Lee, Y.-H. Yang, C.-T. Lee, and H. H. Chen, "Automatic highlights extraction for drama video using music emotion and human face features," *Neurocomputing*, vol. 119, pp. 111–117, Nov. 2013.

[45] C.-T. Lee, Y.-H. Yang and H. H. Chen, "Multipitch estimation of piano music by exemplar-based sparse representation," *IEEE Transactions on Multimedia* (TMM), vol. 14, no. 3, pp. 608–618, Jun. 2012.

[46] Y.-H. Yang and H. H. Chen, "Machine recognition of music emotion: a review," *ACM Transactions on Intelligent Systems and Technology* (TIST), vol. 3, no. 3, article 40, May 2012.

[47] Y.-C. Lin, Y.-H. Yang, and H. H. Chen, "Exploiting online tags for music emotion classification," *ACM Transactions on Multimedia Computing, Communications, and Applications* (TOMCCAP), vol. 7s, no. 1, article 26, Oct. 2011.

[48] Y.-H. Yang and H. H. Chen, "Prediction of the distribution of perceived music emotions using discrete samples," *IEEE Transactions on Audio, Speech, and Language Processing* (TASLP), vol. 19, no. 7, pp. 2184 -2196, Sep. 2011.

[49] Y.-H. Yang and H. H. Chen, "Ranking-based emotion recognition for music organization and retrieval," *IEEE Transactions on Audio, Speech, and Language Processing* (TASLP), vol. 19, no. 4, pp. 762-774,

May 2011.

[50] Y.-F. Su, Y.-H. Yang, M.-T. Lu, and H. H. Chen, “Smooth control of adaptive media playout for video streaming,” *IEEE Transactions on Multimedia* (TMM), vol. 11, no. 7, pp. 1331–1339, Nov. 2009.

[51] Y.-H. Yang, W.-H. Hsu, and H. H. Chen, “Online reranking via ordinal informative concepts for context fusion in concept detection and video search,” *IEEE Transactions on Circuits and Systems for Video Technology* (TCSVT), vol. 19, no. 12, pp. 1880–1890, Dec. 2009.

[52] Y.-H. Yang, Y.-C. Lin, Y.-F Su, and H. H. Chen, “A regression approach to music emotion recognition,” *IEEE Transactions on Audio, Speech, and Language Processing* (TASLP), vol. 16, no. 2, pp. 448–457, Feb. 2008.

- **Conference Papers**

[53] Wei-Jaw Lee, Hsuan-Yu Yeh, Ting-Yi Hu, Chih-Pin Tan, Fang-Duo Tsai, Yi-Hsuan Yang, “MPEcho: A melody and phoneme-aware generative framework for controllable cover song generation,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2026.

[54] Wei-Han Hsu, Chih-Cheng Chang, Bo-Yu Chen, Li Su, Yi-Hsuan Yang, “Separate-and-detect: Unified drum transcription and stem generation via latent diffusion,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2026.

[55] Yen-Tung Yeh, Chung-Jui Chan, Yun-Ning Hung, and Yi-Hsuan Yang, “Rethinking automatic music mixing as sequential stem blending,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2026.

[56] Yuan-Chiao Cheng, Jui-Te Wu, Brian Chen, Yen-Tung Yeh, Yu-Hua Chen, Yi-Hsuan Yang, “StemFX: Learning Mixing Style Representations via Autoregressive FX Chain Prediction on Source-Separated Stems,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2026.

[57] Chih-Heng Chang, Keng-Seng Ho, Chih-Yu Tsai, Kuan-Lin Chen, Yi-Hsuan Yang, and Jian-Jiun Ding, “AnchorSteer: Self-discovered concept injection for structure-preserving music editing,” in *Proc. ACM Conf. Knowledge Discovery and Data Mining* (KDD), 2026.

[58] Fang-Chih Hsieh, Wei-Jaw Lee, Chun-Ping Wang, Hung-yi Lee, Hao-Wen Dong, Yi-Hsuan Yang, “Academic text-to-music grand challenge: Datasets, baselines, and evaluation methods,” *IEEE Int. Conf. Multimedia and Expo.* (ICME), Grand Challenge Paper, 2026.

[59] Jeng-Yue Liu, Ting-Chao Hsu, Yen-Tung Yeh, Li Su, and Yi-Hsuan Yang, "SynthCloner: Synthesizer preset conversion via factorized codec with ADSR envelope control," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2026.

[60] Ting-Kang Wang, Chih-Pin Tan, and Yi-Hsuan Yang, "Time-shifted token scheduling for symbolic music generation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2026.

[61] Xuanjun Chen, Chia-Yu Hu, I-Ming Lin, Yi-Cheng Lin, I-Hsiang Chiu, You Zhang, Sung-Feng Huang, Yi-Hsuan Yang, Haibin Wu, and Jyh-Shing Roger Jang, "How does instrumental music help SingFake detection?," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2026.

[62] Ting-Chao Hsu and Yi-Hsuan Yang, “Conditional vocal timbral technique conversion via embedding-guided dual attribute modulation,” in *AAAI Workshop on Artificial Emerging AI Technologies for Music* (EAIM), 2026.

[63] Jia-Wei Liao, Pin-Chi Pan, Li-Xuan Peng, Sheng-Ping Yang, Yen-Tung Yeh, Cheng-Fu Chou, Yi-Hsuan Yang, “Zero-shot geometry-aware diffusion guidance for music restoration,” in *NeurIPS*

Workshop on Artificial Intelligence for Music (AI4Music), 2025.

[64] Ping-Yi Chen, Chih-Pin Tan, Yi-Hsuan Yang, “Segment-factorized full-song generation on symbolic piano music,” in *NeurIPS Workshop on Artificial Intelligence for Music* (AI4Music), 2025.

[65] Hsin Ai and Yi-Hsuan Yang, “Transformer-based unpaired piano accompaniment style transfer,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf*. (APSIPA ASC), 2025.

[66] Yen-Tung Yeh, Junghyun Koo, Marco Martínez-Ramírez, Wei-Hsiang Liao, Yi-Hsuan Yang, and Yuki Mitsufuji, “Fx-Encoder++: Extracting instrument-wise audio effect representations from mixtures,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2025.

[67] Fang-Duo Tsai, Shih-Lun Wu, Weijaw Lee, Sheng-Ping Yang, Bo-Rui Chen, Hao-Chung Cheng, and Yi-Hsuan Yang, “MuseControlLite: Multifunctional music generation with lightweight conditioners,” in *Proc. Int. Conf. Machine Learning* (ICML), 2025.

[68] Dinh-Viet-Toan Le and Yi-Hsuan Yang, “METEOR: Melody-aware texture-controllable symbolic music re-orchestration via Transformer VAE,” in *Proc. Int. Joint Co*nf. *Artificial Intelligence* (IJCAI), 2025.

[69] Yen-Tung Yeh, Yu-Hua Chen, Yuan-Chiao Cheng, Jui-Te Wu, Jun-Jie Fu, Yi-Fan Yeh, and Yi-Hsuan Yang, "DDSP Guitar Amp: Interpretable guitar amplifier modeling," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2025.

[70] Chon In Leong, I-Ling Chung, Kin Fong Chao, Jun-You Wang, Yi-Hsuan Yang, and Jyh-Shing Roger Jan, “Music2Fail: Transfer music to failed recorder style,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf.* (APSIPA ASC), 2024.

[71] Jingyue Huang, Ke Chen, and Yi-Hsuan Yang, “Emotion-driven piano music generation via two-stage disentanglement and functional representation,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2024.

[72] Yun-Han Lan, Wen-Yi Hsiao, Hao-Chung Cheng, Yi-Hsuan Yang, “MusiConGen: Rhythm and chord control for Transformer-based text-to-music generation,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2024.

[73] Fang-Duo Tsai, Shih-Lun Wu, Haven Kim, Bo-Yu Chen, Hao-Chung Cheng, and Yi-Hsuan Yang, “Audio Prompt Adapter: Unleashing music editing abilities for text-to-music with lightweight finetuning,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2024.

[74] Chih-Pin Tan, Hsin Ai, Yi-Hsin Chang, Shuen-Huei Guan, and Yi-Hsuan Yang, “PiCoGen2: Piano cover generation with transfer learning approach and weakly aligned data,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2024.

[75] Yu-Hua Chen, Yen-Tung Yeh, Yuan-Chiao Cheng, Jui-Te Wu, Yu-Hsiang Ho, Jyh-Shing Roger Jang, and Yi-Hsuan Yang, “Towards zero-shot amplifier modeling: One-to-many amplifier modeling via tone embedding control,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2024.

[76] Yen-Tung Yeh, Wen-Yi Hsiao and Yi-Hsuan Yang, “Hyper recurrent neural network: Condition mechanisms for black-box audio effect modeling,” *in Proc. Int. Conf. Digital Audio Effects* (DAFx), 2024.

[77] Yu-Hua Chen, Woosung Choi, Wei-Hsiang Liao, Marco Martínez-Ramírez, Kin Wai Cheuk, Yuki Mitsufuji, Jyh-Shing Roger Jang and Yi-Hsuan Yang, “Improving unsupervised clean-to-rendered guitar tone transformation using GANs and integrated unaligned clean data,” *in Proc. Int. Conf. Digital Audio Effects* (DAFx), 2024.

[78] Ying-Shuo Lee, Yueh-Po Peng, Jui-Te Wu, Ming Cheng, Li Su and Yi-Hsuan Yang, “Distortion recovery: A two-stage method for guitar effect removal,” *in Proc. Int. Conf. Digital Audio Effects* (DAFx), 2024.

[79] Chih-Pin Tan, Shuen-Huei Guan, Yi-Hsuan Yang, “PiCoGen: Generate piano covers with a two-stage approach,” in *Proc. ACM Int. Conf. Multimedia Retrieval* (ICMR), 2024.

[80] Shih-Lun Wu and Yi-Hsuan Yang, “Compose & Embellish: Well-structured piano performance generation via a two-stage approach,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2023.

[81] Yen-Tung Yeh, Bo-Yu Chen, and Yi-Hsuan Yang, “Exploiting pre-trained feature networks for generative adversarial networks in audio-domain loop generation,” in *Proc. Int. Society for Music Information Retrieval* Conf. (ISMIR), 2022.

[82] Yueh-Kao Wu, Ching-Yu Chiu, and Yi-Hsuan Yang, “Conditional beat-aware drum accompaniment generation in the audio domain using Transformer VQ-VAE,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2022.

[83] Chih-Pin Tan, Wen-Yu Su, and Yi-Hsuan Yang, “Melody infilling with user-provided structural context,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2022.

[84] Da-Yi Wu, Wen-Yi Hsiao, Fu-Rong Yang, Oscar Friedman, Warren Jackson, Scott Bruzenak, Yi-Wen Liu, and Yi-Hsuan Yang, “SawSing: A DDSP-based singing vocoder via subtractive sawtooth waveform synthesis,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2022.

[85] Taejun Kim, Yi-Hsuan Yang, and Juhan Nam, “Joint estimation of fader and equalizer gains of DJ mixers using convex optimization,” *in Proc. Int. Conf. Digital Audio Effects* (DAFx), 2022.

[86] Yu-Chih Tsai, Tse-Yu Pan, Ting-Yang Kao, Yi-Hsuan Yang, and Min-Chun Hu, “EMVGAN: Emotion-aware music-video common representation learning via generative adversarial networks,” in *Proc. Int. Joint Workshop on Multimedia Artworks Analysis and Attractiveness Computing in Multimedia*, in conjunction with ACM ICMR, 2022.

[87] Chien-Feng Liao, Jen-Yu Liu, and Yi-Hsuan Yang, “KaraSinger: Score-free singing voice synthesis with VQ-VAE using Mel-spectrograms,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2022.

[88] Bo-Yu Chen, Wei-Han Hsu, Wei-Hsiang Liao, Marco A. Martínez Ramírez, Yuki Mitsufuji, and Yi-Hsuan Yang, “Automatic DJ transitions with differentiable audio effects and generative adversarial networks,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2022.

[89] Yu-Hua Chen, Wen-Yi Hsiao, Tsu-Kuang Hsieh, Jyh-Shing Roger Jang, and Yi-Hsuan Yang, “Towards automatic transcription of polyphonic electric guitar music: A new dataset and a multi-loss transformer model,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2022.

[90] Fu-Rong Yang, Yin-Ping Cho, Da-Yi Wu, Yi-Hsuan Yang, Shan-Hung Wu, and Yi-Wen Liu, “Mandarin singing voice synthesis with a phonology-based duration model,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf.* (APSIPA ASC), 2021.

[91] Chin-Jui Chang, Chun-Yi Lee, and Yi-Hsuan Yang, “Variable-length music score infilling via XLNet and musically specialized positional encoding,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2021.

[92] Tun-Min Hung, Bo-Yu Chen, Yen-Tung Yeh, and Yi-Hsuan Yang, “A benchmarking initiative for audio-domain music generation using the FreeSound Loop Dataset,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2021.

[93] Hsiao-Tzu Hung, Joann Ching, Seungheon Doh, Nabin Kim, Juhan Nam and Yi-Hsuan Yang, “EMOPIA: A multi-modal pop piano dataset for emotion recognition and emotion-based music generation,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2021.

[94] Juan Gómez-Cañón, Estefania Cano, Yi-Hsuan Yang, Perfecto Herrera, and Emilia Gomez, “Let's agree to disagree: Consensus entropy active learning for personalized music emotion recognition,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2021.

[95] Pedro Sarmento, Adarsh Kumar, C. J. Carr, Zack Zukowski, Mathieu Barthet, and Yi-Hsuan Yang, “DadaGP: A dataset of tokenized GuitarPro songs for sequence models,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2021.

[96] Antoine Liutkus, Ondřej Cífka, Shih-Lun Wu, Umut Simsekli, Yi-Hsuan Yang, and Gael Richard, "Relative positional encoding for Transformers with linear complexity," in *Proc. International Conference on Machine Learning* (ICML), 2021.

[97] Ching-Yu Chiu, Joann Ching, Wen-Yi Hsiao, Yu-Hua Chen, Alvin Wen-Yu Su, Yi-Hsuan Yang, “Source separation-based data augmentation for improved joint beat and downbeat tracking,” in *Proc. European Signal Processing Conference* (EUSIPCO), 2021.

[98] Taejun Kim, Yi-Hsuan Yang, and Juhan Nam, “Reverse-engineering the transition regions of real-world DJ mixes usingsub-band analysis with convex optimization,” in *Proc. International Conference on New Interface for Musical Expression* (NIME), 2021.

[99] Wen-Yi Hsiao, Jen-Yu Liu, Yin-Cheng Yeh, and Yi-Hsuan Yang, “Compound Word Transformer: Learning to compose full-song music over dynamic directed hypergraphs,” in *Proc. AAAI Conf. Artificial Intelligence* (AAAI), 2021 (acceptance rate 21%).

[100] Joann Ching, Antonio Ramires, and Yi-Hsuan Yang, "Instrument role classification: Auto-tagging for loop based music," in *Proc. Joint Conference on AI Music Creativity*, 2020.

[101] Yu-Hua Chen, Yu-Siang Huang, Wen-Yi Hsiao, and Yi-Hsuan Yang, "Automatic composition of guitar tabs by Transformers and groove modeling," in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2020.

[102] Shih-Lun Wu and Yi-Hsuan Yang, " The Jazz Transformer on the front line: Exploring the shortcomings of AI-composed music through quantitative measures," in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2020.

[103] Bo-Yu Chen, Jordan Smith, and Yi-Hsuan Yang, "Neural loop combiner: Neural network models for assessing the compatibility of loops," in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2020.

[104] Taejun Kim, Minsuk Choi, Evan Sacks, Yi-Hsuan Yang, and Juhan Nam, "A computational analysis of real-world DJ mixes using mix-to-track subsequence alignment," in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2020.

[105] Antonio Ramires, Frederic Font, Dmitry Bogdanov, Jordan Smith, Yi-Hsuan Yang, Joann Ching, Bo-Yu Chen, Yueh-Kao Wu, Hsu Wei-Han, and Xavier Serra, "The Freesound Loop Dataset and annotation tool," in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2020.

[106] Jen-Yu Liu, Yu-Hua Chen, Yin-Cheng Yeh and Yi-Hsuan Yang, “Unconditional audio generation with generative adversarial networks and cycle regularization,” in *Proc. INTERSPEECH*, 2020.

[107] Da-Yi Wu and Yi-Hsuan Yang, “Speech-to-singing conversion based on boundary equilibrium GAN,” in *Proc. INTERSPEECH*, 2020.

[108] Ching-Yu Chiu, Wen-Yi Hsiao, Yin-Cheng Yeh, Yi-Hsuan Yang, and Alvin W. Y. Su, “Mixing-specific data augmentation techniques for improved blind violin/piano source separation,” in *Proc. IEEE Int. Workshop on Multimedia Signal Processing* (*MMSP*), 2020.

[109] Yu-Siang Huang and Yi-Hsuan Yang, “Pop Music Transformer: Beat-based modeling and generation of expressive Pop piano compositions,” in *Proc. ACM Int. Conf. Multimedia* (MM), 2020.

[110] Jen-Yu Liu, Yu-Hua Chen, Yin-Cheng Yeh, and Yi-Hsuan Yang, “Score and lyrics-free singing voice generation,” in *Proc. Int. Conf. Computational Creativity* (ICCC), 2020.

[111] Tsung-Han Hsieh, Kai-Hsiang Cheng, Zhe-Cheng Fan, Yu-Ching Yang, Yi-Hsuan Yang, “Addressing the confounds of accompaniments in singer identification,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2020.

[112] Jayneel Parekh, Preeti Rao, Yi-Hsuan Yang, “Speech-to-singing conversion in an encoder-decoder framework,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2020.

[113] Jianyu Fan, Yi-Hsuan Yang, Kui Dong, Philippe Pasquier, “A comparative study of Western and Chinese classical music based on soundscape models,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2020.

[114] Hsiao-Tzu Hung, Chung-Yang Wang, Yi-Hsuan Yang, Hsin-Min Wang, “Improving automatic Jazz melody generation by transfer learning techniques,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf.* (APSIPA ASC), 2019.

[115] Eva Zangerle, Michael Vötter, Ramona Huber and Yi-Hsuan Yang, “Hit song prediction: Leveraging low- and high-level audio features,” in *Proc. Int. Society for Music Information Retrieval Conf*. (ISMIR), 2019.

[116] Kai-Hsiang Cheng, Szu-Yu Chou, and Yi-Hsuan Yang, “Multi-label few-shot learning for sound event recognition,” in *Proc. IEEE Int. Workshop on Multimedia Signal Processing* (*MMSP*), 2019.

[117] Frederic Tamagnan and Yi-Hsuan Yang, “Drum fills detection and generation,” in *Proc. Int. Symp. Computer Music Multidisciplinary Research* (*CMMR*), 2019.

[118] Jen-Yu Liu and Yi-Hsuan Yang, “Dilated convolution with dilated GRU for music source separation,” in *Proc. Int. Joint Co*nf. *Artificial Intelligence* (IJCAI), 2019.

[119] Yun-Ning Hung, I-Tung Chiang, Yi-An Chen, Yi-Hsuan Yang, “Musical composition style transfer via disentangled timbre representations,” in *Proc. Int. Joint Conf. Artificial Intelligence* (IJCAI), 2019.

[120] Zhe-Cheng Fan, Tak-Shing Chan, Yi-Hsuan Yang and Jyh-Shing Jang, “Deep cyclic group networks,” in *Proc. Int. Joint Conf. Neural Networks* (IJCNN), 2019.

[121] Yun-Ning Hung, Yian Chen and Yi-Hsuan Yang, “Multitask learning for frame-level instrument recognition,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2019.

[122] Tsung-Han Hsieh, Li Su, and Yi-Hsuan Yang, “A streamlined encoder/decoder architecture for melody extraction,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2019.

[123] Szu-Yu Chou, Kai-Hsiang Cheng, Jyh-Shing Roger Jang, and Yi-Hsuan Yang, “Learning to match transient sound events using attentional similarity for few-shot sound recognition,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2019.

[124] Chih-Ming Chen, Chuan-Ju Wang, Ming-Feng Tsai and Yi-Hsuan Yang, “Collaborative similarity embedding for recommender systems,” in *Proc. the Web Conference* (WWW), 2019, short paper (acceptance rate 20%).

[125] Vibert Thio, Hao-Min Liu, Yin-Cheng Yeh, and Yi-Hsuan Yang, “A minimal template for interactive web-based demonstrations of musical machine learning,” in *Proc. Workshop on Intelligent Music Interfaces for Listening and Creation* (MILC), 2019.

[126] Bryan Wang and Yi-Hsuan Yang, “PerformanceNet: Score-to-audio music generation with multi-band convolutional residual network,” in *Proc. AAAI Conf. Artificial Intelligence* (AAAI), 2019 (acceptance rate 16%).

[127] Jen-Yu Liu and and Yi-Hsuan Yang, “Denoising auto-encoder with recurrent skip connections and residual regression for music source separation,” in *Proc. IEEE Int. Conf. Machine Learning and Applications* (ICMLA), 2018.

[128] Hao-Ming Liu and and Yi-Hsuan Yang, “Lead sheet generation and arrangement by conditional generative adversarial network,” in *Proc. IEEE Int. Conf. Machine Learning and Applications* (ICMLA), 2018.

[129] Yi-Wen Chen, Yi-Hsuan Yang, and Homer H. Chen, “Cross-cultural music emotion recognition by adversarial discriminative domain adaptation,” in *Proc. IEEE Int. Conf. Machine Learning and Applications* (ICMLA), 2018.

[130] Chia-An Yu, Ching-Lun Tai, Tak-Shing Chan and Yi-Hsuan Yang, “Modeling multi-way relations with hypergraph embedding,” in *Proc. ACM Int. Conf. Information and Knowledge Management* (CIKM), 2018.

[131] Hao-Wen Dong and Yi-Hsuan Yang, “Convolutional generative adversarial networks with binary neurons for polyphonic music generation,” in *Proc. Int. Society for Music Information Retrieval Conf*. (ISMIR), 2018.

[132] Yun-Ning Hung and Yi-Hsuan Yang, “Frame-level instrument recognition by timbre and pitch,” in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2018.

[133] Szu-Yu Chou, Jyh-Shing Roger Jang, and Yi-Hsuan Yang, “Learning to recognize transient sound events using attentional supervision,” in *Proc. Int. Joint Conf. Artificial Intelligence* (IJCAI), 2018.

[134] C.-W. Wu, J.-Y. Liu, Y.-H. Yang, J.-S. R. Jang, “Singing style transfer using cycle-consistent boundary equilibrium generative adversarial networks,” in *Proc. Joint Workshop on Machine Learning for Music*, extended abstract, 2018.

[135] A. Poddar, E. Zangerle, and Y.-H. Yang, “#nowplaying-RS: A new benchmark dataset for building context-aware music recommender systems,” in *Proc. Sound and* Music *Computing Conf.* (SMC), 2018.

[136] W.-L. Wei, J.-C. Lin, T.-L. Liu, Y.-H. Yang, H.-M. Wang, H.-R. Tyan, and H.-Y. M. Liao, “SeetheVoice: Learning from music to visual storytelling of shots,” *Proc. IEEE Int. Conf. Multimedia and Expo.* (ICME), 2018.

[137] H.-W. Dong, W.-Y. Hsiao, L.-C. Yang and Y.-H. Yang, “MuseGAN: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment,” in *Proc. AAAI Conf. Artificial Intelligence* (AAAI), 2018 (acceptance rate 25%).

[138] Y.-S. Huang, S.-Y. Chou and Y.-H. Yang, “Generating music medleys via playing music puzzle games,” in *Proc. AAAI Conf. Artificial Intelligence* (AAAI), 2018 (acceptance rate 25%).

[139] Y.-S. Huang, S.-Y. Chou and Y.-H. Yang, “Music thumbnailing via neural attention modeling of music emotion,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf.* (APSIPA ASC), 2017.

[140] C.-A. Yu, T.-S. Chan and Y.-H. Yang, “Low-rank matrix completion over finite Abelian group algebras

for context-aware recommendation," in *Proc. ACM Int. Conf. Information and Knowledge Management* (CIKM), 2017.

[141] L.-C. Yang, S.-Y. Chou, Y.-H. Yang, "MidiNet: A convolutional generative adversarial network for symbolic-domain music generation," in *Proc. Int. Society for Music Information Retrieval Conf.* (ISMIR), 2017.

[142] C.-C. Shih, P.-C. Li, Y.-J. Lin, A. W. Y. Su, L. Su and Y.-H. Yang, "Analysis and synthesis of the violin playing styles of Heifetz and Oistrakh," *in Proc. Int. Conf. Digital Audio Effects* (DAFx), 2017.

[143] S.-Y. Chou, L.-C. Yang, Y.-H. Yang, and J.-S. Jang, "Conditional preference nets for user and item cold start problems in music recommendation," *Proc. IEEE Int. Conf. Multimedia and Expo.* (ICME), 2017, pp. 1147-1152.

[144] Z.-C. Fan, T.-S. T. Chan, Y.-H. Yang, and J.-S. R. Jang, "Music signal processing using vector product neural networks," *Proc. International Workshop on Deep Learning for Music* (DLM), 2017.

[145] L.-C. Yang, S.-Y. Chou, J.-Y. Liu, Y.-H. Yang, and Yi-An Chen, "Revisiting the problem of audio-based hit song prediction using convolutional neural networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), pp. 621-625, 2017.

[146] T.-W. Su, J.-Y. Liu, and Y.-H. Yang, "Weakly-supervised audio event detection using event-specific Gaussian filters and fully convolutional networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), pp. 791-795, 2017.

[147] S.-Y. Su, C.-K. Chiu, L. Su, and Y.-H. Yang, "Automatic conversion of pop music into chiptunes for 8-bit pixel art," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), pp. 411-415, 2017.

[148] L. Gao, L. Su, Y.-H. Yang, and T. Lee, "Polyphonic piano note transcription with non-negative matrix factorization of differential spectrogram," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), pp. 291-295, 2017.

[149] W.-L. Wei, J.-C. Lin, T.-L. Liu, Y.-H. Yang, H.-M. Wang, H.-R. Tyan, H.-Y. M. Liao, "Deep-net fusion to classify shots in concert videos," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), pp. 1383-1387, 201.

[150] M.-H. Yang, L. Su and Y.-H. Yang, "Highlighting root notes in chord recognition using cepstral features and multi-task learning," in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf*. (*APSIPA ASC*), 2016.

[151] J.-Y. Liu and Y.-H. Yang, "Event localization in music auto-tagging," *in Proc. ACM Multimedia* (*MM*), full paper (acceptance rate 20%), pp. 1048-1057, 2016.

[152] S.-Y. Chou, Y.-H. Yang, J.-S. Jang and Y.-C. Lin, "Addressing cold start for next-song recommendation," *in Proc. ACM Recommender Systems* (*RecSys*), short paper, pp. 115-118, 2016.

[153] C.-M. Chen, M.-F. Tsai, Y.-C. Lin and Y.-H. Yang, "Query-based music recommendations via preference embedding," *in Proc. ACM Recommender Systems* (*RecSys*), short paper, pp. 79-82, 2016.

[154] C.-H. Yang, P.-C. Li, A. W. Y. Su, L. Su, and Y.-H. Yang, "Automatic violin synthesis using expressive musical term features," *in Proc. Int. Conf. Digital Audio Effects* (*DAFx*), 2016.

[155] L. Su, T.-Y. Chuang and Y.-H. Yang, "Exploiting frequency, periodicity and harmonicity using advanced time-frequency concentration techniques for multipitch estimation of choir and symphony," *in Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 393-399, 2016.

[156] Y.-P. Chen, L. Su and Y.-H. Yang, “Electric guitar playing technique detection in real-world recording based on F0 sequence pattern recognition,” *in Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 708-714, 2015.

[157] C.-Y. Liang, L. Su, H.-M. Lin and Y.-H. Yang, “Musical offset detection of pitched instruments: the case of violin,” *in Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 281-287, 2015.

[158] P.-C. Li, L. Su, Y.-H. Yang and A. W. Y. Su, “Analysis of expressive musical terms in violin using score-informed and expression-based audio features,” *in Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 809-815, 2015.

[159] Y.-J. Luo, L. Su, Y.-H. Yang and T.-S. Chi, “Detection of common mistakes in novice violin playing,” *in Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 316-322, 2015.

[160] A. Aljanaki, Y.-H. Yang, and M. Soleymani, “Emotion in Music Task at MediaEval 2015,” in *Proc. MediaEval Workshop*, extended abstract, 2015.

[161] P.-K. Jao, Y.-P. Lin, Y.-H. Yang, and T.-P. Jung, “Using robust principal component analysis to alleviate day-to-day variability in EEG based emotion classification,” in *Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 570-573, 2015.

[162] C.-M. Chen, P.-C. Chien, M.-F. Tsai, Y.-H. Yang and Y.-C. Lin, “Exploiting latent social listening representations for music recommendations,” in *Proc. ACM Recommender Systems* (*RecSys*), poster paper, 2-page poster paper, 2015.

[163] P.-K. Jao, P.-I. Chen, and Y.-H. Yang, “Disk Jockey in Brain - A prototype for volume control of tracked instrument during playback,” in *Proc. Int. Works. Brain-Computer Music Interfacing* (*BCMI*), 2015.

[164] L. Su and Y.-H. Yang, “Escaping from the abyss of manual annotation: New methodology of building polyphonic datasets for automatic music transcription,” in *Proc. Int. Symp. Computer Music Multidisciplinary Research* (*CMMR*), 2015.

[165] P.-I. Chen, J.-Y. Liu, and Y.-H. Yang, “Personal factors in music preference and similarity: User study on the role of personality traits,” in *Proc. Int. Symp. Computer Music Multidisciplinary Research* (*CMMR*), 2015.

[166] S.-Y. Chou, Y.-H. Yang, and Y.-C. Lin, “Evaluating music recommendation in a real-world setting: On data splitting and evaluation metrics,” in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), 2015.

[167] P.-K. Jao, Y.-H. Yang, and B. Wohlberg, “Informed monaural source separation of music based on convolutional sparse coding,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2015.

[168] T.-S. Chan, T.-C. Yeh, Z.-C. Fan, H.-W. Chen, L. Su, Y.-H. Yang, and J.-S. Jang, “Vocal activity informed singing voice separation with the IKALA dataset,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2015.

[169] Y.-A. Chen, Y.-H. Yang, J.-C. Wang, H. H. Chen, “The AMG1608 dataset for music emotion recognition,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2015.

[170] C.-H. Yeh, Y.-H. Yang, M.-H. Chang, and H.-Y. M. Liao, “Music driven human motion manipulation for characters in a video,” in *Proc. IEEE Int. Symp. Multimedia* (*ISM*), 2014.

[171] C.-M. Chen, H.-P. Chen, M.-F. Tsai, and Y.-H. Yang, “Leverage item popularity and recommendation quality via cost-sensitive factorization machines,” in *Proc. IEEE Int. Conf. Data Mining* (*ICDM*), Ph.D. forum paper, 2014.

[172] C.-C. M. Yeh, P.-K. Jao, and Y.-H. Yang, “AWtoolbox: Characterizing audio information using audio words,” in *Proc. ACM Multimedia* (*MM*), short paper, 2014.

[173] M. Soleymani, A. Aljanaki, Y.-H. Yang, M. N. Caro, F. Eyben, K. Markov, B. Schuller, R. Veltkamp, F. Weninger, and F. Wiering, “Emotional analysis of music: A comparison of methods,” in *Proc. ACM Multimedia* (*MM*), short paper, 2014.

[174] A. Aljanaki, Y.-H. Yang, and M. Soleymani, “Emotion in Music Task at MediaEval 2014,” in *Proc. MediaEval Workshop*, extended abstract, 2014.

[175] L. Su, L.-F. Yu, and Y.-H. Yang, “Sparse cepstral and phase codes for guitar playing technique classification,” *in Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), 2014.

[176] J.-C. Wang, M.-C. Yan, Y.-H. Yang and H.-M. Wang, “Automatic set list identification and song segmentation of full-length concert videos,” *in Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), 2014.

[177] L. Su and Y.-H. Yang, “Power-scaled spectral flux and peak-valley group-delay methods for robust musical onset detection,” in *Proc*. *Sound and* Music *Computing Conf.* (*SMC*), 2014.

[178] L. Su, L.-F. Yu, Y.-H. Yang, and Hsin-Yu Lai, “Resolving octave ambiguities: A cross-dataset Investigation,” in *Proc*. *Sound and* Music *Computing Conf.* (*SMC*), 2014.

[179] X. Hu and Y.-H. Yang, “A study on cross-cultural and cross-dataset generalizability of music mood regression models,” in *Proc*. *Sound and* Music *Computing Conf.* (*SMC*), pp. 1149–1155, 2014.

[180] X. Hu and Y.-H. Yang, “Cross-cultural mood regression for music digital libraries,” in *Proc*. *IEEE/ACM Joint Conf. Digital Libraries* (*DL*), 2014. [Best Poster Award]

[181] J.-Y. Liu, S.-Y. Liu and Y.-H. Yang, “LJ2M Dataset: Toward better understanding of music listening behavior and user mood,” in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), 2014.

[182] S.-Y. Wang, J.-C. Wang, Y.-H. Yang and H.-M. Wang, “Towards time-varying music auto-tagging based on CAL500 Expansion,” in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), 2014.

[183] P.-K. Jao, C.-C. M. Yeh and Y.-H. Yang, “Modified LASSO screening for audio word-based music classification using large-scale dictionary,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2014.

[184] L.-F. Yu, L. Su and Y.-H. Yang, “Sparse cepstral codes and power scale for instrument identification,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2014.

[185] C.-C. M. Yeh, J.-C. Wang, Y.-H. Yang and H.-M. Wang, “Improving music auto-tagging by intra-song instance bagging,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2014.

[186] Y.-A. Chen, J.-C. Wang, Y.-H. Yang and H. H. Chen, “Linear regression-based adaptation of music emotion recognition models for personalization,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2014.

[187] L. Su and Y.-H. Yang, “Sparse modeling for artist identification: Exploiting phase information and vocal separation," in *Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 349–354, 2013.

[188] Y.-H. Yang, “Low-rank representation of both singing voice and music accompaniment via learned dictionaries,” in *Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 427–432, 2013.

[189] M. Soleymani, M. N. Caro, E. Schmidt, C.-Y. Sha, and Y.-H. Yang, “1000 songs for emotional analysis of music,” in *Proc. Int. Workshop on Crowdsourcing for Multimedia* (*CrowdMM*), pp. 1–6, 2013.

[190] M. Soleymani, M. N. Caro, E. Schmidt, C.-Y. Sha, and Y.-H. Yang, “The MediaEval 2013 brave new task: Emotion in Music,” in *Proc. MediaEval Workshop*, in conjunction with ACM Multimedia, extended abstract, 2013.

[191] C.-M. Chen, M.-F. Tsai, J.-Y. Liu, and Y.-H. Yang, “Using Emotional Context from Article for Contextual Music Recommendation,” in *Proc. ACM Multimedia* (*MM*), short paper, 2013.

[192] C.-M. Chen, M.-F. Tsai, J.-Y. Liu, and Y.-H. Yang, “Music recommendation based on multiple contextual similarity information,” in *Proc. IEEE/WIC/ACM Int. Conf. Web Intelligence* (*WI*), 2013.

[193] P.-K. Jao, L. Su, and Y.-H. Yang, “Analyzing the dictionary properties and sparsity constraints for a dictionary-based music genre classification system,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf*. (*APSIPA ASC*), 2013.

[194] C.-C. Yeh and Y.-H. Yang, “Towards a more efficient sparse coding based audio-word feature extraction system,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf*. (*APSIPA ASC*), 2013.

[195] Y.-H. Yang, “Towards real-time music auto-tagging using sparse features,” in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), oral paper (best paper finalist), 2013.

[196] Y.-C. Teng, Y.-S. Kuo, and Y.-H. Yang, “A large in-situ data set for context-aware music recommendation on smartphones,” in *Proc. IEEE Int. Conf. Multimedia and Expo*. (*ICME*), short paper, 2013.

[197] C.-C. Yeh, L. Su, and Y.-H. Yang, “Dual-layer bag-of-frames model for music genre classification,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing* (*ICASSP*), pp. 246–250, 2013.

[198] C.-Y. Sha, Y.-H. Yang, and H. H. Chen, “Singing voice timbre classification of Chinese popular music,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing* (*ICASSP*), pp. 734–738, 2013.

[199] J.-C. Wang, Y.-H. Yang, H.-M. Wang, and S.-K. Jeng, “Personalized music emotion recognition via model adaptation,” in *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conf*. (*APSIPA ASC*), 2012.

[200] J.-C. Wang, Y.-H. Yang, H.-M. Wang, and S.-K. Jeng, “The Acoustic Emotion Gaussians model for emotion-based music annotation and retrieval,” in *Proc. ACM Multimedia* (*MM*), full paper (accept rate=20%), pp. 89–98, 2012.

[201] J.-Y. Liu, C.-C. Yeh, Y.-C. Teng, and Y.-H. Yang, “Bilingual analysis of song lyrics and audio words,” in *Proc. ACM Multimedia* (*MM*), short paper, pp. 829–832, 2012.

[202] Y.-H. Yang, “On sparse and low-rank matrix decomposition for singing voice separation,” *Proc. ACM Multimedia* (*MM*), short paper, pp. 757–760, 2012.

[203] J.-Y. Liu and Y.-H. Yang, “Inferring personal traits from music listening history,” in *Proc. Int. Workshop on Music Information Retrieval with User-Centered and Multimodal Strategies (MIRUM),* pp. 31–36, 2012.

[204] J.-C. Wang, Y.-H. Yang, K.-C. Chang, H.-M. Wang, and S.-K. Jeng, “Exploring the relationship between categorical and dimensional emotion semantics of music,” in *Proc. Int. Workshop on Music Information Retrieval with User-Centered and Multimodal Strategies (MIRUM),* pp. 63–68, 2012.

[205] Y.-H. Yang and X. Hu, “Cross-cultural music mood classification: A comparison of English and Chinese songs,” in *Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 19–24, 2012.

[206] C.-C. Yeh and Y.-H. Yang, “Supervised dictionary learning for music genre classification,” in *Proc. ACM*

Int. Conf. Multimedia Retrieval (ICMR), 2012.

[207] Y.-H. Yang, D. Bogdanov, P. Herrera, and M. Sordo, "Music retagging using label propagation and robust principal component analysis," in *Proc. Int. Workshop on Advances in Music Information Research* (*AdMIRe*), pp. 869–876, 2012.

[208] C.-D. Lee, Y.-H. Yang, and H. H. Chen, "Automatic transcription of piano music by sparse representation of magnitude spectra," in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), oral paper, 2011.

[209] Y.-H. Kuo, H.-T. Lin, W.-H. Cheng, Y.-H. Yang, and W.-H. Hsu, "Unsupervised auxiliary visual words discovery for large-scale image object retrieval," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition* (*CVPR*), 2011.

[210] K.-S. Lin, A. Lee, Y.-H. Yang, C.-T. Lee and H. H. Chen, "Automatic highlights extraction for drama video using music emotion and human face features," in *Proc. IEEE Int. Workshop on Multimedia Signal Processing* (*MMSP*), 2011.

[211] Y.-H. Yang, Y.-C. Lin, and H. H. Chen, "Personalized music emotion recognition," in *Proc. ACM Int. Conf. Information Retrieval* (*SIGIR*), pp. 748–749, 2009.

[212] Y.-H. Yang, Y.-C. Lin, A. Lee, and H. H. Chen, "Improving musical concept detection by ordinal regression and context fusion," in *Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 147–152, 2009.

[213] M.-Y. Su, Y.-H. Yang, Y.-C. Lin, and H. H. Chen, "An integrated approach to music boundary detection," in *Proc. Int. Society for Music Information Retrieval Conf.* (*ISMIR*), pp. 705–710, 2009.

[214] Y.-H. Yang and H. H. Chen, "Music emotion ranking," in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing* (*ICASSP*), pp. 1657–1660, 2009.

[215] Y.-H. Yang, Y.-C. Lin, and H. H. Chen, "Clustering for music search results," in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), pp. 874–877, 2009.

[216] Y.-C. Lin, Y.-H. Yang, and H. H. Chen, "Exploiting genre for music emotion classification," in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), pp. 618–621, 2009.

[217] H.-T. Cheng, Y.-H. Yang, Y.-C. Lin, and H. H. Chen, "Multimodal structure segmentation and analysis of music using audio and textual information," in *Proc. IEEE Int. Symp. Circuits and Systems* (*ISCAS*), pp. 1677–1680, 2009.

[218] L.-C. Hsieh, K.-T. Chen, C.-H. Chiang, Y.-H. Yang, G.-L. Wu, C.-S. Ferng, H.-W. Hsueh, C.-R. Tsia, and W.-H. Hsu, "Canonical image selection and efficient image graph construction for large-scale Flickr photos," *Grand Challenge for ACM Multimedia* (*MM*), pp. 1121–1122, 2009.

[219] Y.-H. Yang, P.-T. Wu, C.-W. Lee, K.-H. Lin, W.-H. Hsu, and H. H. Chen, "ContextSeer: Context search and recommendation at query time for shared consumer photos," *Proc. ACM Multimedia* (*MM*), full paper (accept rate=20%), pp. 199–208, 2008.

[220] P.-T. Wu, Y.-H. Yang, and W.-H. Hsu, "Keyword-based concept search on consumer photos by web-based kernel function," in *Proc. ACM Multimedia* (*MM*), short paper, pp. 651–654, 2008.

[221] Y.-H. Yang and W.-H. Hsu, "Video search reranking via online ordinal reranking," *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), 2008.

[222] Y.-H. Yang, Y.-C. Lin, H.-T. Cheng, I.-B. Liao, Y.-C. Ho, and H. H. Chen, "Toward multi-modal music emotion classification," in *Proc. Pacific-Rim Conf. Multimedia* (*PCM*), pp. 70–79, 2008.

[223] H.-T. Cheng, Y.-H. Yang, Y.-C. Lin, I.-B. Liao, and H. H. Chen, "Automatic chord recognition for music

classification and retrieval," in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), pp. 1505–1508, 2008.

[224] Y.-H. Yang, Y.-F. Su, Y.-C. Lin, and H. H. Chen, "Music emotion recognition: The role of individuality," in *Proc. Int. Workshop on Human-centered Multimedia* (*HCM*), pp. 13–21, 2007.

[225] Y.-H. Yang, Y.-C. Lin, Y.-F Su, and H. H. Chen, "Music emotion classification: A regression approach," in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), pp. 208–211, 2007.

[226] M.-F. Weng, C.-K. Chen, Y.-H. Yang, R.-E. Fan, Y.-T. Hsieh, Y.-Y. Chunag, W.-H. Hsu, and C.-J. Lin, "The NTU toolkit and framework for high-level feature detection at TRECVID 2007," *NIST TRECVID Workshop* 2007.

[227] C.-C. Ma, Y.-H. Yang, and W. Hsu, "Image thumbnailing via multi-view face detection and saliency analysis," in *Proc. Conf. Visual Information Processing* (*VIP*), *20*07.

[228] Y.-H. Yang, C.-C Liu, and H. H. Chen, "Music emotion classification: A fuzzy approach," in *Proc. ACM Multimedia* (*MM*), short paper, pp. 81–84, 2006.

[229] C.-C. Liu, Y.-H. Yang, P.-H. Wu, and H. H. Chen, "Detecting and classifying emotion in popular music," in *Proc. Joint Int. Conf. Information Sciences & Int. Conf. Computer Vision, Pattern Recognition and Image Processing 2006* (*JCIS/CVPRIP*), pp. 996–999, 2006.

[230] Y.-H. Yang, M.-T Lu, and H. H. Chen, "Smooth playout control for video streaming over error-prone channels," in *Proc. IEEE Int. Symp. Multimedia* (*ISM*), pp. 415–418, 2006.

- **Technical Report**

[231] Yuan-Chiao Cheng, Yi-Hsuan Yang, Vincent K.M. Cheung, "ProCraft: Instruction-driven music production from free-form text," in ArXiv e-prints, abs/2611.07423, November 2026.

[232] Hsueh-Wei Fu, Yen-Tung Yeh, and Yi-Hsuan Yang, "Expressive guitar synthesis via controllable latent flow matching," in ArXiv e-prints, abs/2610.36786, October 2026.

[233] Wei-Han Hsu, Li-Jie Lin, Po-Hsuan Lai, Po-Hsiang Huang, Jeng-Yue Liu, Li Su, and Yi-Hsuan Yang, "DJustify: Reason-and-Verify DJ planning for Pop transitions and story-driven sets," in ArXiv e-prints, abs/2610.33434, October 2026.

[234] Yun-Chen Cheng, Chih-Pin Tan, Tzu-Hung Huang, and Yi-Hsuan Yang, "Piano2Pop: Expanding piano demos into Pop productions without aligned supervision," in ArXiv e-prints, abs/2610.31510, October 2026.

[235] Chi-En Dai, Chih-Pin Tan, and Yi-Hsuan Yang, "Predicting music valence and arousal via textual and multimodal in-context learning in large language models," in ArXiv e-prints, abs/2610.31202, October 2026.

[236] Wei-Jaw Lee, Chun-Ping Wang, Xuanjun Chen, and Yi-Hsuan Yang, "MQAScore: Fine-grained text-music alignment evaluation via large audio language model," in ArXiv e-prints, abs/2610.31083, October 2026.

[237] Kuan-Hsun Ho, Yuan-Chiao Cheng, Yen-Tung Yeh, and Yi-Hsuan Yang, "MixJudge: Learning a language-queryable critic for vocal mixing," in ArXiv e-prints, abs/2610.30783, October 2026.

[238] Song-Ze Yu, Fang-Duo Tsai, and Yi-Hsuan Yang, "MIDI-REPA: Improving text-to-music generation with symbolic representation alignment," in ArXiv e-prints, abs/2610.20967, October 2026.

[239] Yen-Tung Yeh, Yun-Ning Hung, and Yi-Hsuan Yang, "EG-VAE: A unified framework for electric guitar

tone transfer and removal," in ArXiv e-prints, abs/2608.05513, August 2026.

[240] Fang-Duo Tsai, Yi-An Lai, Fei-Yueh Chen, Hsueh-Wei Fu, Li Chai, Hao-Chung Cheng, and Yi-Hsuan Yang, "MIDI-informed singing accompaniment generation in a compositional song pipeline," in ArXiv e-prints, abs/2602.22029, February 2026.

[241] Wei-Jaw Lee, Fang-Chih Hsieh, Xuanjun Chen, Fang-Duo Tsai, and Yi-Hsuan Yang, "Training-efficient text-to-music generation with state-space modeling," in ArXiv e-prints, abs/2601.14786, January 2026.

[242] Jingyue Huang and Yi-Hsuan Yang, "Emotion-driven melody harmonization via melodic variation and functional representation," in ArXiv e-prints, abs/2407.20176, July 2024.

[243] Wei-Han Hsu, Bo-Yu Chen, and Yi-Hsuan Yang, "Deep learning based EDM subgenre classification using Mel-spectrogram and tempogram features," in ArXiv e-prints, abs/2110.08862, submitted October 2021.

[244] H.-W. Dong and Y.-H. Yang, "Towards a deeper understanding of adversarial losses," in: ArXiv e-prints, abs/1901.08753, submitted January 2019.

[245] Y.-N. Hung, Y.-A. Chen and Y.-H. Yang, "Learning disentangled representations for timber and pitch in music audio," in: ArXiv e-prints, abs/1811.03271, submitted November 2018.

[246] H.-W. Dong and Y.-H. Yang, "Training generative adversarial networks with binary neurons by end-to-end backpropagation," in: ArXiv e-prints, abs/1810.04714, submitted October 2018.

[247] C.-M. Chen, Y.-H. Yang, Y.-A. Chen, M.-F. Tsai, "Vertex-context sampling for weighted network embedding," in: ArXiv e-prints, abs/1711.00227, submitted November 2017.

[248] L.-C. Yu, Y.-H. Yang, Y.-N. Hung, Y.-A. Chen, "Hit song prediction for pop music by Siamese CNN with ranking loss," in: ArXiv e-prints, abs/1710.10814, submitted October 2017.

[249] J.-Y. Liu, S.-K. Jeng and Y.-H. Yang, "Applying topological persistence in convolutional neural network for music audio signals," in: ArXiv e-prints, abs/1608.07373, submitted August 2016.

[250] C.-C. M. Yeh, P.-K. Jao, and Y.-H. Yang, "The AWtoolbox for Characterizing Audio Information," Technical report, TR-CITI-15-001, Academia Sinica, March 2015.

- **Extended abstracts**

[251] Ping Wang, Fang-Duo Tsai, Wei-Jaw Lee, and Yi-Hsuan Yang, "SheetSage-Pro: Bilingual lyric-aligned lead-sheet transcription from music audio," *Int. Society for Music Information Retrieval Conf.*, late-breaking and demo paper (ISMIR-LBD), 2026.

[252] Fang-Duo Tsai and Yi-Hsuan Yang, "Demonstrating singing accompaniment capabilities for MuseControlLite," in *NeurIPS Workshop on Artificial Intelligence for Music* (AI4Music), demo paper, 2025.

[253] Wei-Jaw Lee, Fang-Chih Hsieh, Xuanjun Chen, Fang-Duo Tsai, and Yi-Hsuan Yang, "Exploring state-space-model based language model in music generation," *Int. Society for Music Information Retrieval Conf.*, late-breaking and demo paper (ISMIR-LBD), 2025.

[254] Yen-Tung Yeh, Chung-Jui Chan, and Yi-Hsuan Yang, "DiffFx: A toolkit for differentiable audio effects processors," *Int. Society for Music Information Retrieval Conf.*, late-breaking and demo paper (ISMIR-LBD), 2025.

[255] Yen-Tung Yeh, Wen-Yi Hsiao, and Yi-Hsuan Yang, "PyNeuralFx: A Python package for neural audio effect modeling," *Int. Society for Music Information Retrieval Conf.*, late-breaking and demo paper

(ISMIR-LBD), 2024.

[256] Yu-Hua Chen, Yuan-Chiao Cheng, Yen-Tung Yeh, Jui-Te Wu, Yu-Hsiang Ho, Jyh-Shing Roger Jang, and Yi-Hsuan Yang, "Demo of zero-shot guitar amplifier modelling: Enhancing modeling with hyper neural networks," *Int. Society for Music Information Retrieval Conf.*, late-breaking and demo paper (ISMIR-LBD), 2024.

[257] Yu-Hua Chen, Woosung Choi, WeiHsiang Liao, Marco A. Martinez Ramirez, Kin Wai Cheuk, Yi-Hsuan Yang, Yuki Mitsufuji, “Neural amplifier modelling with several GAN variants,” *Int. Society for Music Information Retrieval Conf*., late-breaking and demo paper (ISMIR-LBD), 2023.

[258] Chih-Pin Tan, Chin-Jui Chang, Alvin W. Y. Su, and Yi-Hsuan Yang, "Music score expansion with variable-length infilling," *Int. Society for Music Information Retrieval Conf*., late-breaking and demo paper (ISMIR-LBD), 2021.

[259] Joann Ching and Yi-Hsuan Yang, "Learning to generate piano music with sustain pedals," *Int. Society for Music Information Retrieval Conf*., late-breaking and demo paper (ISMIR-LBD), 2021.

[260] Hsiao-Tzu Hung, Yu-Hua Chen, Maximilian Mayerl, Michael Vötter, Eva Zangerle, Yi-Hsuan Yang, “MediaEval 2019 emotion and theme recognition task: A VQ-VAE based approch,” MediaEval working note paper, 2019.

[261] Maximilian Mayerl, Michael Vötter, Hsiao-Tzu Hung, Bo-Yu Chen, Yi-Hsuan Yang, Eva Zangerle, “Recognizing song mood and theme using convolutional recurrent neural networks,” MediaEval working note paper, 2019.

[262] Wen-Yi Hsiao, Yin-Cheng Yeh, Yu-Siang Huang, Chung-Yang Wang, Jen-Yu Liu, Tsu-Kuang Hsieh, Hsiao-Tzu Hung, Jun-Yuan Wang, and Yi-Hsuan Yang, “Jamming with Yating: Interactive demonstration of a music composition AI,” *Int. Society for Music Information Retrieval Conf*., late-breaking and demo paper (ISMIR-LBD), 2019.

[263] Yin-Cheng Yeh, Jen-Yu Liu, Wen-Yi Hsiao, Yu-Siang Huang, and Yi-Hsuan Yang, “Learning to generate Jazz and Pop piano music from audio via MIR techniques,” *Int. Society for Music Information Retrieval Conf*., late-breaking and demo paper (ISMIR-LBD), 2019.

[264] Y.-H. Chen, B. Wang, and Y.-H. Yang, “Demonstration of PerformanceNet: A convolutional neural network model for score-to-audio music generation,” in *Proc. Int. Joint Co*nf. *Artificial Intelligence* (IJCAI), demo paper, 2019.

[265] H.-W. Dong, W.-Y. Hsiao, and Y.-H. Yang, "Pypianoroll: Open source Python package for handling multitrack pianorolls," *Int. Society for Music Information Retrieval Conf*., late-breaking and demo paper (ISMIR-LBD), 2018.

[266] H.-M. Liu, M.-H. Wu, and Y.-H. Yang, "Lead sheet generation and arrangement via a hybrid generative model," *Int. Society for Music Information Retrieval Conf*., late-breaking and demo paper (ISMIR-LBD), 2018.

[267] S.-Y. Chou, J.-S. R Jang, and Y.-H. Yang, "FrameCNN: A weakly-supervised learning framework for frame-wise acoustic event detection and classification," *Proc. Detection and Classification of Acoustic Scenes and Events Workshop* (DCASE), extended abstract, 2017.

[268] H.-W. Dong, W.-Y. Hsiao, L.-C. Yang, and Y.-H. Yang, "MuseGAN: Demonstration of a convolutional GAN based model for generating multi-track piano-rolls," *Proc. Int. Society for Music Information Retrieval Conf*., demo paper (ISMIR-LBD), 2017.

[269] Y.-S. Huang, S.-Y. Chou, and Y.-H. Yang, "DJnet: A Dream for Making An Automatic DJ," *Int. Society*

for Music Information Retrieval Conf., late-breaking and demo paper (ISMIR-LBD), 2017.

[270] J.-W. Peng, S.-W. Sun, W.-H. Cheng, and Y.-H. Yang, “eMosic: Mobile media pushing through social emotion sensing,” in *Proc. ACM Multimedia* (*MM*), demo paper, 2015.

[271] A. Aljanaki, Y.-H. Yang, and M. Soleymani “Musical emotion variation detection from acoustic content - lessons learned from developing MediaEval "Emotion in Music" benchmark,” in *Proc. Int. Conf. Music and Emotion* (ICME4), 2015.

[272] W.-C. Lin, S.-W. Sun, W.-. Cheng, Y.-H. Yang, K.-L. Hua, F.-J. Wang, and J.-J. Wang, “Attaching-Music: an interactive music delivery system for private listening as wherever you go,” in *Proc. IEEE Int. Conf. Multimedia and Expo.* (*ICME*), demo paper, 2014.

[273] J.-C. Wang, Y.-H. Yang, H.-M. Wang, “The acousticvisual emotion Guassians model for automatic generation of music video,” in *Proc. ACM Multimedia* (*MM*), grand challenge solution paper (First Prize), pp. 1379–1380, 2012.

[274] Y.-H. Kuo, Y.-L. Wu, K.-T. Chen, Y.-H. Yang, T.-H. Chiu, and W.-H. Hsu, “A technical demonstration of large-scale image object retrieval by efficient query evaluation and effective auxiliary visual feature discovery,” in *Proc. ACM Multimedia* (*MM*), technical demo, 2010.

[275] T.-L. Wu, H.-K. Wang, C.-C. Ho, Y.-P. Lin, T.-T. Hu, M.-F. Weng, L.-W. Chan, C.-H. Yang, Y.-H. Yang, Y.-P. Hung, Y.-Y. Chuang, H.-H. Chen, H. H. Chen, J.-H. Chen, and S.-K. Jeng, “Interactive content presention based on expressed emotion and physiological feedback,” in *Proc. ACM Multimedia* (*MM*), technical demonstration, pp. 1009–1010, 2008.

[276] Y.-H. Yang, Y.-C. Lin, H.-T. Cheng, and H. H. Chen, “Mr. Emo: Music retrieval in the emotion plane,” in *Proc. ACM Multimedia* (*MM*), technical demo, pp. 1003–1004, 2008.

- **Book Chapters**

[277] J.-C. Wang, Y.-H. Yang and H.-M. Wang, “Affective music information retrieval,” In *Emotions and Personality in Personalized Services*, M. Tkalčič *et al.* Eds., Springer, 2016.

[278] Y.-H. Yang, J.-C. Wang, Y.-A. Chen, and H. H. Chen, “Model Adaptation for Personalized Music Emotion Recognition,” In *Handbook of Pattern Recognition and Computer Vision*, C.-H. Chen Eds., World Scientific Publishing, Feb. 2016.

- **Dissertation**

[279] Y.-H. Yang, *Dimensional Music Emotion Recognition for Content Retrieval*, defended at Jan. 21, 2010.

PATENTS

- H. H. Chen and Y.-H. Yang, “Search device and associated methods,” USA patent, US 8,666,910 B2, Mar. 4, 2014.
- Y.-H. Yang and H. H. Chen, “Digital data processing method for personalized information retrieval and computer readable storage medium and information retrieval system thereof,” USA patent, US 8,321,412 B2, Nov. 27, 2012.
- Y.-H. Yang and H. H. Chen, “用於模擬個體差異之個人化資訊檢索之數位資料處理方法及其電腦裝置可讀式資訊儲存媒體與資訊檢索系統 (Digital data processing method for personalized information retrieval and computer readable storage medium and information retrieval system thereof),” Taiwan patent, TW I396105, May 2013.

- W.-L. Chang, I.-B. Liao, Y.-C. Lin, Y.-H. Yang, and H. H. Chen, “基於曲風之音樂情緒二層式分類法 (A genre-based two-layer structure for music emotion classification),” Taiwan patent, TW I380285, Dec. 21, 2012.
- Y.-H. Yang and H. H. Chen, “多媒體檢索系統、建立該系統的方法及其應用方法 (A Multimedia Searching System, a Method of Building the System and Associate Searching Method Thereof),” Taiwan patent, TW 201022968 Jun. 2010.

REFERENCES

Available upon request