CHAEWAN CHUN

czc5884@psu.edu | website

KEYWORDS: Machine Learning; Generative AI; Large Language Models; Computer Vision; Speech & Audio Processing; Multimodal Learning; Information Retrieval

EDUCATION

The Pennsylvania State University *University Park, PA* Aug 2022 – Present
Ph.D. candidate in **Informatics** | Advisor: Dr. Dongwon Lee GPA: 3.94/4.0

The Pennsylvania State University, *University Park, PA* Aug 2018 – May 2022
Schreyer Honors College, Dean's List (every semester) GPA: 3.91/4.0
B.S. in **Computer Science**, *Magna Cum Laude* | B.S. in **Mathematics**, *Cum Laude* | Minor: **Statistics**
HONORS THESIS: Robust Image Classification based on Pixel Importance (Advisor: Dr. Jia Li).

RESEARCH EXPERIENCES

Benchmarking Fact-Checking in Spoken Media Jan 2025 – Present
Pennsylvania State University

- Built the MAD → MAD2 benchmarks for spoken misinformation, scaling to thousands of turns with two-speaker audio, turn-aligned transcripts, and claim-level check-worthiness & veracity labels.
- Leveraged LLMs end-to-end: generated multi-turn dialogues from seed claims and performed batch linguistic screening to ensure naturalness and coverage; defined tasks and reference baselines at sentence- and dialogue-level.

Audio Fact-Checking Pipelines & Context Modeling Jan 2025 – Present
Pennsylvania State University

- Designed an end-to-end verification stack for conversational audio (claim detection → retrieval → verification → explanation) and systematically evaluated temporal context regimes (claim-only, past-only/causal, $\pm K$-second windows, full dialogue).
- Implemented a gated audio–text verifier (WavLM for speech; DeBERTa-v3 for transcripts) with shared-space projections, consistently outperforming audio-only across context settings.

Large Vision-based Video Clip Dataset with Emotion Labels May 2023 – Dec 2024
Pennsylvania State University

- Co-developed a large-scale video emotion corpus with psychology/kinesiology collaborators; built pose-based time-series clustering over joint keypoints to characterize motion dynamics.
- Created an iterative, transfer-learning data-collection pipeline to mitigate label imbalance and improve robustness.

Assorting Claude Monet's Water Lily Paintings via Computer Vision Aug 2022 – Dec 2024
Pennsylvania State University

- Implemented brushstroke *flow* extraction from cropped flower patches using structure-tensor fields and streamline tracing to model bristle-mark directionality and flow coherence; computed turning angles/curvature features for downstream analysis.
- Supported style/chronology analyses by supplying distributional, stroke-level features (curvature statistics, wiggliness, segment lengths) used by the project's classifiers.

Robust Image Classification based on Pixel Importance May 2021 – May 2022
Pennsylvania State University

- Proposed *heatmap-guided* training: compute explanation maps to flag low-importance pixels, retain the

top ∼15%, blur the remainder (Gaussian σ=4), and train on these selectively blurred images to reduce overfitting and encourage robustness.
- Evaluated under adversarial attacks (FGSM, NewtonFool): observed consistent robustness gains over standard training, while noting an adaptive "attack-after-blur" threat model can erode accuracy (contrast: perturb-before-blur shows smaller drops).

Publications

Chun, Chaewan, Delvin Ce Zhang, and Dongwon Lee. "When Misinformation Speaks and Converses: Rethinking Fact-Checking in Audio Platforms." *ACL Rolling Review (ARR)*, **under review**, 2026.

Chun, Chaewan, Delvin Ce Zhang, and Dongwon Lee. "Listening for Lies: Context-Sensitive Claim Verification in Spoken Dialogues." *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, **under review**, 2026.

Meruyert Aristombayeva, Jason Lucas, **Chaewan Chun**, and Dongwon Lee. "Detecting Spoken Hallucinations Across Three Languages: Generation, Challenges, and Insights from Audio Data." *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).* **under review**, 2026.

Chun, Chaewan, Lysandre Terrisse, Delvin Ce Zhang, and Dongwon Lee. "MAD: A Benchmark for Multi-Turn Audio Dialogue Fact-Checking." *International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS)*, accepted as working paper, 2025.

Li, Jia, **Chaewan Chun**, Kathryn Brown, and James Z. Wang. "Computational Investigation of Abstraction in Claude Monet's Water Lilies through Brushstroke Analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **under review**, 2025.

Awards & Grants

Fund for Excellence in Graduate Recruitment Award, *academic achievement, Penn State*	Jun 2022
Paul MacDonald Open Doors Scholarship, *leadership and academic achievement, Penn State*	Nov 2021
Leadership Scholarship, *leadership engagement in Women in Engineering, Penn State*	May 2021
Student Engagement Network Innovation Grant, *summer research, Penn State*	Apr 2021
Evan Pugh Scholar Award, *academic achievement, Penn State*	Apr 2020

Others

Leadership:	Teaching Assistant, Interdisciplinary Research Design for IST (IST 501) | Learning Assistant, Programming Language Concepts (CMPSC 461) | Administrative Assistant - data analytics | Engineering Orientation Network Leader/Mentor | Penn State Peer Mentor Collective | Facilitated Study Group Coordinator, Women in Engineering.
Posters:	MASC–SLL Symposium 2025 | SBP–BRiMS 2025.