Tongda Xu

x.tongda@nyu.edu | 185-1020-1763 | Github: https://github.com/tongdaxu
Apt 1507, 62 Xinwenhua St, Xicheng Dist, Beijing

I am a PhD candidate in Tsinghua University, advised by Dr. Ya-Qin Zhang. My research interests include generative modeling and neural compression. I have published 5 papers as first author in NIPS, ICML and ICLR (including 1 spotlight).

Education

Tsinghua University — Beijing, China
Doctor of Philosophy in Computer Science & Technology, Advised by Dr. Ya-Qin Zhang — *Sep 2023 -*

- **Relevant Coursework**: Applied Stochastic Process, Information Theory, Advanced Theoretical Computer Science

New York University — New York, US
Master of Science in Digital Media, GPA: 3.85/4.0, Advised by Dr. Wang Yao — *Sep 2018 - Dec 2020*

- **Awards**: Graduate Fellowship, Graduate Research Fellowship
- **Relevant Coursework**: Image and Video Processing, Machine Learning, Probability and Stochastic Process, Optimization and Numerical Linear Algebra, Data Structure and Algorithm, Computer System Architecture
- **Thesis**: Accelerating Crowd Volunteer Audio Description Production with Content Recommendation

Tsinghua University — Beijing, China
Bachelor of Architecture, GPA: 3.82/4.0, Advised by Dr. Huang Weixin — *Aug 2013 - May 2018*

- **Awards**: National Fellowship, Tsinghua-Nitianzeng Fellowship, Academic Excellence Fellowship ×2
- **Relevant Coursework**: Calculus, Computer Graphics
- **Thesis**: Mediating People's EEG Alpha Band Power through Reinforcement Learning

Experience

NVIDIA Corporation — Beijing, China
AI4Science Intern, Advised by Zhixiang Dai — *Nov 2024 - Feb 2025*

- **Estimating Sample Uncertainty in Diffusion based AI4Science model**: Fast sample variance estimation for diffusion based surface solar radiation downwards prediction. Project will be presented in GTC 2025 by Jensen Huang.

Microsoft Research Asia (MSRA) — Beijing, China
Research Intern (Media Computing Group), Excellent Intern, Advised by Dr. Li Bin — *Feb 2024 - Nov 2024*

- **Pre-trained Diffusion Model as a Perceptual Constraint**: Using pre-trained diffusion model as a GAN discriminator.

Institute for AI Industry Research (AIR), Tsinghua University — Beijing, China
Research Assistant (Neural Codec), Advised by Dr. Wang Yan — *June 2022 - Present*

- **Bit-allocation using Optimization** : Prove the equivalence between optimal bit allocation for video codec via semi-amortized variational inference. Published in ICML 2023.

Sensetime Research — Beijing, China
Researcher (Codec Group), Advised by Dr. Wang Yan — *Jan 2021 - June 2022*

- **Semi-amortized Inference for Variable Bitrate Image Compression**: Propose an approach to achieves variable bitrate model and perception-distortion trade-off in neural image compression using semi-amortized variational inference. Published in NeurIPS 2022 [c6].
- **Importance Weighted Neural Image Compression**: We propose an approach for neural image compression training with tighter ELBO. Published in NeurIPS 2022 [c5].
- **Spatial Moment Pooling for Blind Image Quality Assessment**: Propose to extent spatial average pooling into spatial moment pooling by incorporating higher order moments. Published in ICIP 2022 [c4].
- **Compressed Domain Contour Flow for Fast Video Salient Detection**: Propose a novel contour-flow approach for P/B frame salient object detection by wrapping the I frame contours with motion vectors encoded in bitstream. The boundary problem of feature-flow approaches is overcame by wrapping contours directly. Achieve 400% speed up with 3% f-measure loss compared with SOTA.

Microsoft Research Asia (MSRA) — Beijing, China
Research Intern (Media Computing Group), Excellent Intern, Advised by Dr. Li Bin — *Jan 2020 - Dec 2020*

- **Joint Rate Control for Real-time Communication with Reinforcement Learning**: Propose and implement a deep reinforcement learning based framework for joint rate control of multiple sources in real-time communication. Outperform WebRTC + VP8 in screen and video sharing scenario by 0.5db in PSNR.
- **Fluid Screen Sharing**: Propose and implement a mixed-integer linear programming based algorithm for screen content re-layout. Facilitate real-time relaying out for screen sharing across different devices.
- **Error Recovery for Video Communication**: Dive in, implement and evaluate core video error recovery algorithm including Hybrid Type I, WebRTC and Microsoft in-house RTC.

Video Lab, New York University New York, US

Research Intern (Image Processing), Advised by Dr. Wang Yao *May 2019 - Aug 2019*

- **Optimal Feature Subset Selection for MRI**: Propose a novel improvement of A* heuristic search for feature selection, combining the idea of A* graph search and genetic algorithm, improved state of the art of problem by 0.2 in r2 score for regression.
- **Image Cascade Network for Large 3D Medical Image Segmentation**: Propose and implement an end-to-end auto-context based multi-level network structure enabling semantic segmentation for 3D image as large as 256^3. Achieve dice coefficient 0.99 for background, 0.92 for mouse embryo and 0.85 for brain vehicle.

Aedas Ltd (HK Office) Hong Kong, China

Architecture Intern, Advised by Christine Lam and David Clayton *May 2017 - Aug 2017*

- **Computer-Aided Design Toolkit**: Optimize working pipeline and accelerate design, calculation and analysis process from feasibility study to construction for 6 complex buildings in China. Compile a computational toolkit with C# in Grasshopper, facilitating building codes calculation and parametric curtain wall design.

X-Studio, Tsinghua University Beijing, China

Research Intern (Human-Computer Interaction), Advised by Dr. Mi Haipeng *May 2015 - Aug 2015*

- **Tangible Circuit Training**: Aid design and implement of a table based tangible user interface application simulating circuits and electric components. Aid graphic and interface programming in Java.

TEACHING

DM-UY 1143 - Ideation & Prototyping, New York University New York, US

Teaching Assistant to Ideation and Prototyping of Benedetta Piantella *Sep 2018 - Jan 2019*

Architecture Design Studio 6 - Parametric Design , Tsinghua University New York, US

Teaching Assistant to Parametric Design studio of Dr. Weixin Huang *May 2015 - Aug 2015*

PUBLICATIONS

- **[c10]**: Zhang, X., Ge, X., **Xu, T.**, He, D., Wang, Y., Qin, H., ... & Zhang, J. (2025). Gaussianimage: 1000 fps image representation and compression by 2d gaussian splatting. In European Conference on Computer Vision (ECCV).
- **[c9]**: **Xu, T.**, Zhu, Z., He, D., Li, Y., Guo, L., Wang, Y., ... & Zhang, Y. Q. (2024). Idempotence and Perceptual Image Compression. The Twelfth International Conference on Learning Representations (ICLR).
- **[c8]**: Li, Y., **Xu, T.**, Wang, Y., Liu, J., & Zhang, Y. Q. (2023, November). Idempotent Learned Image Compression with Right-Inverse. In Thirty-seventh Conference on Neural Information Processing Systems (NIPS).
- **[c7]**: **Xu, T.***, Gao, H.*, Gao, C., Pi, J., Li, Y., Wang, Y., ... & Wang, Y. (2023). Bit allocation using optimization. International Conference on Machine Learning. (ICML) (*equal contribution)
- **[c6]**: Gao, C.*, **Xu, T***., He, D., Wang, Y., Qin, H. (2022). Flexible Neural Image Compression via Code Editing. Advances in Neural Information Processing Systems. (NIPS) (*equal contribution)
- **[c5]**: **Xu, T.**, Wang, Y., He, D., Gao, C., Gao, H., Liu, K. Qin, H. (2022). Multiple-sample Neural Image Compression. Advances in Neural Information Processing Systems. (NIPS)
- **[c4]**: **Xu, T.**, Shao, Y., Wang, Y.,& Qin, H. (2022). Spatial Moment Pooling Improves Neural Image Assessment. In 2021 IEEE International Conference on Image Processing (ICIP). IEEE.
- **[c3]**: He, D., Yang, Z., Yu, H., **Xu, T.**, Luo, J., Chen, Y., ... & Wang, Y. (2022). PO-ELIC: Perception-Oriented Efficient Learned Image Coding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop (pp. 1764-1769).
- **[j1]**: Qiu, Z., **Xu, T.**, Langerman, J., Das, W., Wang, C., Nair, N., ... & Wang, Y. (2021). A Deep Learning Approach for Segmentation, Classification and Visualization of 3D High Frequency Ultrasound Images of Mouse Embryos. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control.
- **[c2]**: **Xu, T.***, Qiu, Z.*, Das, W., Wang, C., Langerman, J., Nair, N., ... & Wang, Y. (2020, April). Deep Mouse: An End-to-End Auto-Context Refinement Framework for Brain Ventricle & Body Segmentation in Embryonic Mice Ultrasound Volumes. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (pp. 122-126). IEEE.(*equal contribution)
- **[c1]**: **Xu, T.**, Wang, D., & You, X. (2018, October). Mindgame: Mediating People's EEG Alpha Band Power through Reinforcement Learning. In The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings (pp. 5-6).

COMMUNITY SERVICE

- **Program Committee / Reviewer**: NIPS 23-24, ICML 25, ICLR 24-25, CVPR 23-25, ICCV & ECCV 23-24, AISTATS 25, etc.

PATENTS

- **Granted**: CN113660531B, CN113612999B
- **Published**: US20240195968A1, CN114782385A, CN114240787A,CN115695800A, CN116437087A, CN116112671A, CN116600112A