Yingtao Tian — Research Statement

Computational approach with creativity with humanities.

Figure 1: Our work [20] generates abstract artwork for text prompt "Self."

Creativity has consistently been benefiting from the technology innovations [14] including recent advances of machine learning — for example powerful generative models for text, image, audio and video. However, a wide gap still exists for important creative works besides the high-profile applications, as those with culture, artists and behavior considerations or focus has yet to get fully boosted from these advances in machine learning. This presents challenges, especially when **designer's discretion on the system's behavior** or **culture consideration** are presented. In my research I aim at combining **computational approach**, with **generative/creative settings** with artists, culture, humanities and designer's consideration, in a way that the considerations are met while machine learning can help boosting the performances. To do so, I propose techniques and tools that have both **addressed the needs in creative settings** and **advances in core machine learning**. They include 1) generating artifacts with designer's discretion, 2) machine learning boosted tools for historical and cultural data, and 3) the advancing machine learning techniques and tools concerning evolution strategy, optimal transport, language and graph learning.

Figure 2: Our work [19] generates abstract artwork coherently satisfying two text prompts "Walt Disney World." and "a daytime picture of Tokyo."

Generating Artifacts with Artistical Discretion

Besides high-profile application of generative models, a wide range of creative tasks that require designer's artistic discretion lacks full potential from machine learning. I identify that a co-design of generating algorithm and generating algorithm is needed to exercise such discretion. In my work ES-CLIP [20] and ES-3D [18] I **design one particular abstract art** form, framing as synthesizing painting by placing semi-transparent triangles respectively on 2D canvas (Figure 1 and 3D space. To optimize such work I co-design a parameterization of this process as a fitting function for **Evolutionary Strategy**, enabling a unique style as a result of the co-design.

Allowing larger controllability with existing methods is also an important problem in enabling designer's discretion. In my work Simultaneous Multiple-Prompt [19] I designed a fully end-to-end differential approach that can make any text-to-image models adapting to multiple text prompts coherently. For example it generates images coherently with multiple text prompts (Figure 2), tackled by combining **text-to-image generative models** with **optimal transport**.

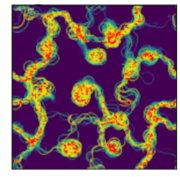

Figure 3: Our work [15] produces collective agent behaviours through environment-enabled communication.

Recently I also look at artifact generation beyond visual models, and broadly to the learned behavior that one can control through designing the environment. In our work Evolving Collective AI [15] we bridge

biological collective intelligence with artificial intelligence, by constructing **designs of agents** inspired by the biological ants collectively solving tasks while using chemical pheromone for communication. It leads to interesting dynamics where collective intelligence is realized in our AI models. We also extend this line of work in a few ALife research communities.

Machine learning-Boosted Tools for Historical and Cultural works

Figure 4: Our work [21, 23] proposes dataset for historical Japanese artworks.

If we walk a little bit away from purely creative settings where creation is mostly boundless, we can encounter a lot of works related to historical and cultural aspects where faithful adherence to a few paradigms is a must. Among them there are many fields with such traits, which poses a challenge when creative, generative models are being applied to them, for which I identify a few cases among East Asia historical art and culture, and propose solutions as follows.

The first line of works resides in Japanese historical art. My work KaoKore [21] and ARC Ukiyo-e Face [23] (Figure 4) organizes **facial expression images** cropped from 14-16 century and 17-19 century Japanese artworks respectively. With learned metadata, they not only preserve culture data but also enable **computational analysis** to such artworks. For example, learned features identify artists with unique qualities.

Recently, we follow up by our work MingOfficial [6] where we process LLMs and GNNs to learn representations for Chinese Ming Dynasty officials from structured data and raw historical record, enabling humanity researchers to quickly **identify historical figures** with interesting traits impossible before. We also, in our work Digital Typhoon [9], organize the longest consecutive typhoon satellite image dataset in modern history to enable ML models helping **weather prediction**.

Furthermore, broadly for CJK (Chinese, Japanese, Korean) topics,I focused on combining generative models and culture-rich context. By designing proper diffusion models, we propose DiffCJK [17] to **generate glyphs of new calligraphy and typology styles** for hundreds of thousands CJK characters. It could generate each CJK glyph's calligraphy style based on printed forms. This leads to an open source model and collaboration for the font community.

Advancing Machine Learning Techniques

As the advances in the creative models cannot be without advances in core machine learning research and I conduct a wide range of research in this direction with my colleagues and student researchers I mentor.

A great amount of work concerns **evolution strategy**, an optimization technique other than the more commonly used gradient based methods. We propose the EvoJAX [16], a scalable, general purpose, hardware-accelerated neuroevolution toolkit; NeuroEvoBench [10], a new benchmark of evolutionary optimization methods tailored toward Deep Learning applications and exhaustively evaluating traditional and meta-learned tasks. We further extend evolution strategy with LLM in a few works: EvoLLM [11] successfully implements a type of black-box recombination operator using LLM through Evolution, be the first showing the potential of LLMs in the evolution setting, and EvoTransformer [12] flexibly characterizes a family of evolution strategies with Transformer architecture, exhibit exhibits for the first time strong optimization performance otherwise challenging neuroevolution tasks

Besides evolution strategy, some work also concerns **reinforcement learning**: In our work DEIR [24], we propose a theoretically-backend intrinsic reward from conditional mutual information to quickly learn policies through faster exploration. We also build OTT [7], an **optimal transport** tool implemented with the JAX framework, to get auto-diff, parallel and jit-able computations. Some of my works during PhD focus on the representation learning side of computational approach to **graphs** [1, 22, 8] and **knowledge bases** [3, 13, 2, 5, 4] , which are well-cited for providing early grounds in these fields.

Future Research Directions

My research interest heavily places on both the **core machine learning** and its creativity application in cases where special considerations, such as appreciation for **art, culture, or collective and emerging behavior**, are desired. This will include three broad directions that I am most excited about.

Emergent behaviors from designing systems with large complexity. My use of the evolution algorithm, a less-common optimization algorithm, in abstract art creation demonstrates that the complexity of the system could lead to unique, emergent behaviors. Moving forward, I would seek to explore sophisticated emerging behaviors from increased complexity in models and mechanisms. First, complexity can result from larger models other than the simple neural network parameterization of agents. Here recent advances in diffusions and LLMs/transformers, as my result core machine learning works suggests, can be a good candidate.

Secondly, by looking at complexity in multi-agent settings, I am also looking for enabling complex emergent behaviors like collaboration, division of labor, or competition, and one way could be working on artificial life where agents follow biological or physical inspired mechanisms in their

interactions.

Computational creativity to fields with art and culture appreciations. The abstract art creation (ES-CLIP, ES-3D), and the research paradigm I explored for computational approaches to East Asian arts and histories (Kaokore, ARC Ukiyo-e Face, MingOfficial), are just two kinds of examples where many possible research can help bridge the gap between them. Overall, institutional collections, like many museums, which while doing a good job of digitalization still lacks computational analysis to it. It remains to be seen what we can harvest from such rich collections, especially the ones in European museums.

More broadly, another important area is how we can fundamentally improve the art and culture research paradigm to bring computational approaches *a priori*. I argue that the current status is that computational research is used mostly as convenient tools and this is not enough as many endeavors are wasted in "patching" a machine learning method to constraints it is never meant to address. Based on my previous works, I seek to collaborate with art and culture researchers on designing novel pipelines to fulfill the potential of recent ML advances, and engage in the community through organizing workshops, etc.

I am also interested in going beyond single art form, say images and text, to develop multi-modality art and culture work creation, as improvements on multi-modality generative models accelerate recently. This would include a joint creation of visual and text art, especially in narratives, where consistency and creator's discretion matters while fast interaction loops are needed.

Core machine learning research on optimization and foundation models. My current work on researching evolution strategy and involvement in improving optimal transports make me believe that advancing the computational creativity would nonetheless require working on machine learning models themselves. Such works are not only inspired by actual application but also open doors for more downstreaming works, and more importantly, by not working on the crowd gradient based optimization it enables more possibilities. In the future I seek to continue my research on evolution strategies, including its use as a black-box optimization, both for its own sake and enabling hard-to-optimize systems, should they become necessary. I believe that these less-commonly used optimization tools can lead to different solutions, just like my abstract art work shows, and is often underlooked.

Moreover, diving into more complex models like diffusions and transferformers for emergent behavior means more special and sophisticated algorithms will be needed for a successful and efficient exploration. For instance in the work EvoLLM and EvoTransformer we find LLM can serve

as a well-functioning numerical optimizer, and this means they have the potential of a large model in planning and interaction, which is worthy of diving deeper. I plan to explore new capital from such complex models, for the sake of a complex system, and as always, release them as open-source software and engage with a wide range of collaboration with external stockholders including universities, research institutes, artists and open-source enthusiasts.

References: My Works

[1] Haochen Chen et al. "Fast and accurate network embeddings via very sparse random projection". In: *Proceedings of the 28th ACM international conference on information and knowledge management*. 2019, pp. 399–408.

[2] Muhao Chen et al. "Co-training embeddings of knowledge graphs and entity descriptions for cross-lingual entity alignment". In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-18)*. 2018.

[3] Muhao Chen et al. "Learning to Represent Bilingual Dictionaries". In: *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*. 2019, pp. 152–162.

[4] Muhao Chen et al. "Multilingual knowledge graph embeddings for cross-lingual knowledge alignment". In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-17)*. 2016.

[5] Muhao Chen et al. "On2vec: Embedding-based relation prediction for ontology population". In: *Proceedings of the 2018 SIAM International Conference on Data Mining*. SIAM. 2018, pp. 315–323.

[6] You-Jun Chen et al. "MingOfficial: A Ming Official Career Dataset and a Historical Context-Aware Representation Learning Framework". In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. 2023, pp. 4380–4401.

[7] Marco Cuturi et al. "Optimal transport tools (ott): A jax toolbox for all things wasserstein". In: *arXiv preprint arXiv:2201.12324* (2022).

[8] Hanjun Dai et al. "Syntax-Directed Variational Autoencoder for Structured Data". In: *International Conference on Learning Representations*. 2018.

[9] Asanobu Kitamoto et al. "Digital Typhoon: Long-term Satellite Image Dataset for the Spatio-Temporal Modeling of Tropical Cyclones". In: *Thirty-seventh Conference on Neural Information Processing (NeurIPS 2023) Systems Datasets and Benchmarks Track*. 2023.

[10] Robert Tjarko Lange, Yujin Tang, and Yingtao Tian. "NeuroEvoBench: Benchmarking Evolutionary Optimizers for Deep Learning Applications". In: *Thirty-seventh Conference on Neural Information Processing (NeurIPS 2023) Systems Datasets and Benchmarks Track*. 2023.

[11] Robert Tjarko Lange, Yingtao Tian, and Yujin Tang. "Evolution Transformer: In-Context Evolutionary Optimization". In: *The Genetic and Evolutionary Computation Conference (GECCO) 2024*. 2024.

[12] Robert Tjarko Lange, Yingtao Tian, and Yujin Tang. "Large Language Models As Evolution Strategies". In: *The Genetic and Evolutionary Computation Conference (GECCO) 2024*. 2024.

[13] Jingjing Li et al. "Spatialnli: A spatial domain natural language interface to databases using spatial comprehension". In: *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 2019, pp. 339–348.

[15] Ryosuke Takata et al. "Evolving Collective AI: Simulation of Ants Communicating via Chemicals". In: *ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference*. MIT Press. 2023.

[16] Yujin Tang, Yingtao Tian, and David Ha. "Evojax: Hardware-accelerated neuroevolution". In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. 2022, pp. 308–311.

[17] Yingtao Tian. "DiffCJK: Conditional Diffusion Model for High-Quality and Wide-coverage CJK Character Generation". In: *arXiv preprint arXiv:2404.05212* (2043).

[18] Yingtao Tian. "Evolving Three Dimension (3D) Abstract Art: Fitting Concepts by Language". In: *arXiv preprint arXiv:2304.12932* (2023).

[19] Yingtao Tian, Marco Cuturi, and David Ha. "Simultaneous Multiple-Prompt Guided Generation Using Differentiable Optimal Transport". In: *Proceedings of the twelfth International Conference on Computational Creativity, ICCC'22*. 2022.

[20] Yingtao Tian and David Ha. "Modern Evolution Strategies for Creativity: Fitting Concrete Images and Abstract Concepts". In: *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*. 2022, pp. 275–291.

[21] Yingtao Tian et al. "Kaokore: A pre-modern japanese art facial expression dataset". In: *Proceedings of the twelfth International Conference on Computational Creativity, ICCC'22*. 2020.

[22] Yingtao Tian et al. "Social relation inference via label propagation". In: *Advances in Information Retrieval: 41st European Conference on IR Research, ECIR 2019, Cologne, Germany, April 14–18, 2019, Proceedings, Part I 41*. Springer. 2019, pp. 739–746.

[23] Yingtao Tian et al. "Ukiyo-e analysis and creativity with attribute and geometry annotation". In: *Proceedings of the twelfth International Conference on Computational Creativity, ICCC'21*. 2021.

[24] Shanchuan Wan et al. "DEIR: Efficient and Robust Exploration through Discriminative-Model-Based Episodic Intrinsic Rewards". In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)*. 2023.

Other References

[14] Christiane Paul. *Digital art*. Thames & Hudson, 2023.