Generative Artificial Intelligence in Research and Teaching

> Hsuan-Tien Lin 林軒田

Professor, National Taiwan University



April 24, 2024, ASUS/AICS

# From Intelligence to Artificial Intelligence

intelligence: thinking and acting smartly

- humanly
- rationally

artificial intelligence: computers thinking and acting smartly

- humanly
- rationally

humanly  $\approx$  smartly  $\approx$  rationally —are humans rational?  $\bigcirc$ 

## Humanly versus Rationally

What if your self-driving car decides one death is better than two—and that one is you? (The Washington Post http://wpo.st/ZK-51)

You're humming along in your self-driving car, chatting on your iPhone 37 while the machine navigates on its own. Then a swarm of people appears in the street, right in the path of the oncoming vehicle.

#### Car Acting Humanly

to save my (and passengers') life, stay on track

#### **Car Acting Rationally**

avoid the crowd and crash the owner for minimum total loss

#### which is smarter? —depending on where I am, maybe? ☺

# Traditional vs. Modern [My] Definition of AI

#### **Traditional Definition**

humanly  $\approx$  intelligently  $\approx$  rationally

#### My Definition

 $\begin{array}{l} \text{intelligently} \approx \text{easily} \\ \text{is your smart phone `smart'? } \bigcirc \end{array}$ 

modern artificial intelligence = application intelligence

# Machine Learning Connects (Big) Data and AI From Big Data to Artificial Intelligence big data $\longrightarrow$ ML $\longrightarrow$ artificial (application) intelligence dish

Photos Licensed under CC BY 2.0 from Andrea Goh on Flickr

#### many possibilities when using the right tools

H.-T. Lin (NTU)

## From AI to GAI: Is This GAI? (1/4)



Photos Licensed under CC BY-SA 3.0 from Diacritica on Wikimedia Commons

generative, but arguably no intelligence

# From AI to GAI: Is This GAI? (2/4)



Photos Licensed under CC BY-SA 4.0 from Lovepeacejoy404 on Wikimedia Commons

#### predictive intelligence, but arguably not generative

H.-T. Lin (NTU)

# From AI to GAI: Is This GAI? (3/4)



Leonardo da Vinci, in Public Domain



Van Gogh, in Public Domain all images are downloaded from Wikipedia



Pjfinlay, with CC0

generative intelligence, or just (predictive) image processing?

# Some Expected Properties of Generative AI



Pjfinlay, with CC0

sophisticated output

structurally intelligent

#### some diversity

creatively intelligent

#### or maybe humanoid?

H.-T. Lin (NTU)

# From AI to GAI: Is This GAI? (4/4)



predictive: time-series prediction; generative: sophisticated output; or does it matter? :-)

# A Story on Modern Generative AI



## Neural Network: from Bird to Airplanes



by UC Regents Davis campus-brainmaps.org. Licensed under CC BY 3.0 via Wikimedia Commons





by Lauris Rubenis. Licensed under CC BY 2.0 via https://flic.kr/ p/fkVuZX



by Pedro Ribeiro Simões. Licensed under CC BY 2.0 via https://flic.kr/ p/adiV7b

#### neural network: a bio-inspired model

H.-T. Lin (NTU)

# From Wright Flyer (1903) to Commercial Airplanes (1919–)



by Wright Brothers. Licensed under Public Domain via US Library of Congress



by Pedro Ribeiro Simões. Licensed under CC BY 2.0 via https://flic.kr/ p/adiV7b

#### we are at wright-flyer-age of (generative) AI

H.-T. Lin (NTU)

## What's Needed before Wider Acceptance

- war? :-)
- technology advancements
  - -like lighter materials, more efficient engines, better control
- regulations
  - -like laws, licenses, etc.
- trials



# will start from research on better control, and then discuss some trial(-and-error) attempts

H.-T. Lin (NTU)

# Si-An Chen, Chun-Liang Li, and Hsuan-Tien Lin. A unified view of cGANs with and without classifiers. NeurIPS 2021.

special thanks to AICS for supporting Si-An Chen via Ph.D. program between 2020 and 2023

# **Generative Models**



#### generative models: creativity for machines

# Generative Adversarial Network (GAN)



# high-quality generation when generator can fool the discriminator

# Conditional Generative Adversarial Network (cGAN)



high-quality conditional (i.e. controllable) generation when generator can fool the conditional discriminator

H.-T. Lin (NTU)

# Existing Attempts to Design cGANs



# can we comprehensively understand and improve them?

H.-T. Lin (NTU)

# **Our Contributions**

#### a comprehensive study of cGANs, which ...

- explains the designs of classifier loss, unconditional loss, conditional loss, contrastive loss in principle
- unifies ACGAN, ProjGAN, ContraGAN to a new architecture ECGAN with the principled explanation
- achieves state-of-the-art conditional generation performance with the unification

#### will describe key ideas behind the unification without detailed math

# Idea 1: Connecting Loss Terms

- classifier loss: for estimating  $p(y|\mathbf{x})$
- unconditional loss: for estimating p(x)
- conditional loss: for estimating  $p(\mathbf{x}|y)$

-how do they connect?

#### Bayes Rule

$$\log p(\mathbf{x}, y) = \log p(\mathbf{x}|y) + \underbrace{\log p(y)}_{\text{assumed as known constants}}$$

$$= \log p(y|\mathbf{x}) + \log p(\mathbf{x})$$

unifying view  $\Rightarrow$  estimate joint distribution  $p(\mathbf{x}, y)$  better from different angles

# Idea 2: Energy-Based Parameterization

unifying view

estimate  $p(\mathbf{x}, y)$  from different angles

#### energy-based parameterization

• if  $\exp(h_{\theta}(\mathbf{x}, y)) \propto p(\mathbf{x}, y)$ ,

$$\log p(\mathbf{x}, y) = \underbrace{h_{\theta}(\mathbf{x}, y)}_{-\log(\text{normalization term})}$$

neural network

log(normalization term) ≈ contrastive loss (proof omitted ☺)

ECGAN (Energy-based Conditional GAN): unifying view + energy-based parameterization

## ECGAN



- unifying view: stabilizes estimation by two angles instead of one
- energy-based parameterization: explains contrastive loss as regularization from energy normalization

#### not just ad-hoc combination, but principled design from math

H.-T. Lin (NTU)



#### complete ECGAN:

2-angle with classifier (w/ or w/o contrastive)

H.-T. Lin (NTU)

# **Evaluation Setting**

Dat	a					
		# training	# test	# classes	resolution	
	CIFAR-10	50,000	10,000	10	32x32	
	Tiny ImageNet	100,000	10,000	200	64x64	
	ImageNet	1,281,167	50,000	1,000	128x128	

#### **Evaluation Metrics**

- Frechet Inception Distance (FID) (↓): unconditional generation goodness
- Inception Score (IS) (<sup>+</sup>): unconditional image quality
- intra-FID (per-class FID) ( $\downarrow$ ): conditional generation goodness

#### is complete ECGAN better?

# Comparison to Existing cGANs

Dataset	Backbone	method	$FID(\downarrow)$	IS (†)	Intra-FID $(\downarrow)$
	DCGAN	ACGAN	$32.507 \pm 2.174$	$7.621 \pm 0.088$	$129.603 \pm 1.212$
		ProjGAN	$21.918 \pm 1.580$	$8.095 \pm 0.185$	$68.164 \pm 2.055$
CIFAR-10		ContraGAN	$28.310 \pm 1.761$	$7.637 \pm 0.125$	$153.730 \pm 9.965$
		ECGAN-UC	$\textbf{18.035} \pm 0.788$	$\textbf{8.487} \pm 0.131$	$\textbf{59.343} \pm 1.557$
	ResGAN	ACGAN	$10.073 \pm 0.274$	$9.512 \pm 0.050$	$48.464 \pm 0.716$
		ProjGAN	$10.195 \pm 0.203$	$9.268\pm0.139$	$46.598 \pm 0.070$
		ContraGAN	$10.551 \pm 0.976$	$9.087\pm0.228$	$138.944 \pm 12.582$
		ECGAN-UC	$\textbf{9.244} \pm 0.062$	$\textbf{9.651} \pm 0.098$	$\textbf{43.876} \pm 0.384$
	BigGAN	ACGAN	$8.615 \pm 0.146$	$9.742 \pm 0.041$	$45.243 \pm 0.129$
		ProjGAN	$8.145\pm0.156$	$9.840 \pm 0.080$	$42.110 \pm 0.405$
		ContraGAN	$8.617 \pm 0.671$	$9.679 \pm 0.210$	$114.602 \pm 13.261$
		ECGAN-UC	$\textbf{7.942} \pm 0.041$	$\textbf{10.002} \pm 0.120$	$\textbf{41.425} \pm 0.221$
		ACGAN	$29.528 \pm 4.612$	$12.964 \pm 0.770$	$315.408 \pm 1.171$
Tiny ImageNet	BigGAN	ProjGAN	$28.451 \pm 2.242$	$12.213 \pm 0.624$	$242.332 \pm 11.447$
		ContraGAN	$24.915 \pm 1.222$	$13.445 \pm 0.371$	$257.657 \pm 3.246$
		ECGAN-UC	$\textbf{18.780} \pm 1.291$	$\textbf{17.475} \pm 1.052$	$\textbf{204.830} \pm 5.648$

#### is complete ECGAN (w/o contrastive) better? YES

## State-of-the-art Performance (2021)

Method	$FID(\downarrow)$	IS(†)
BigGAN*	24.68	28.63
ContraGAN*	25.16	25.25
ECGAN-UC	30.05	26.47
ECGAN-UCE	12.16	56.33
ECGAN-UCE (40k step)	8.49	80.69

Table 6: Evaluation on ImageNet128×128. (\*: Reported by StudioGAN.)

(StudioGAN: https://github.com/POSTECH-CVLab/PyTorch-StudioGAN)

deeper mathematical understanding  $\implies$  state-of-the-art performance

## Generated Images from ECGAN



#### high-quality controllable generation somewhat achieved

H.-T. Lin (NTU)

# Half Summary

#### a comprehensive study of cGANs, which ...

- explains the designs of classifier loss, unconditional loss, conditional loss, contrastive loss in principle
  - connecting three loss terms with Bayes rule
  - applying energy-based models to reveal contrastive loss
- unifies ACGAN, ProjGAN, ContraGAN to a new architecture ECGAN with the principled explanation



achieves state-of-the-art conditional generation

performance with the unification

# fundamental research on math principles $\implies$ controllability

H.-T. Lin (NTU)

Paul Kuo-Ming Huang, Si-An Chen, and Hsuan-Tien Lin. Score-based conditional generation with fewer labeled data by self-calibrating classifier guidance. Work-In-Progress 2024.



# high-quality generation when score function can be estimated

H.-T. Lin (NTU)

# **Conditional SGM**

#### SGM

high-quality unconditional generation when  $\nabla_{\mathbf{x}} \log p(\mathbf{x})$  can be estimated

#### **Conditional SGM**

high-quality conditional generation when  $\nabla_{\mathbf{x}} \log p(\mathbf{x}|y)$  can be estimated

#### Hello Again, Bayes Rule

$$abla_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{y}) = \nabla_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{y})$$

$$\log p(\mathbf{x})$$

unconditional score

$$-\underbrace{\nabla_{\mathbf{x}}\log p(\mathbf{y}|\mathbf{x})}_{\mathbf{x}}$$

 $-\nabla_{\mathbf{x}}\log p(\mathbf{y})$ 

classifier gradient

#### simple CGSGM by classifier guidance + unconditional SGM

# Simple CGSGM

$$\nabla_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{y}) = \underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{x})}_{\text{unconditional score}} + \underbrace{\nabla_{\mathbf{x}} \log p(\mathbf{y}|\mathbf{x})}_{\text{classifier gradient}}$$

#### Pros

- easy reuse of well-trained unconditional SGM
- naturally applicable to semi-supervised data (few labeled data)

# Cons overfitting classifier ⇒ bad conditional score ⇒ bad conditional generation

#### but few labeled data $\Rightarrow$ overfitting classifier?!

H.-T. Lin (NTU)

Key Idea: Align Classifier with Unconditional SGM energy-based parameterization  $\exp(h_{\theta}(\mathbf{x}, y)) \propto p(\mathbf{x}, y)$ 

> unconditional SGM: approximate  $\nabla_x \log p(\mathbf{x})$  by



classifier can cosplay as (be cast to) unconditional SGM

can regularize classifier by its unconditional SGM loss to avoid overfitting (proof omitted )

H.-T. Lin (NTU)

# Comparison to Original CGSGM

with merely 5% of labeled data



#### ours: better quality & more accurate

H.-T. Lin (NTU)

# More Summary



- creativity can go wild —regularization by another view helps control
  - ECGAN: another view by decomposing joint probability
  - improved CGSGM: another view by casting classifier as unconditional SGM
- most importantly, math helps! : —more efforts on fundamental research needed
  - Bayes rule helps
  - energy-based parameterization helps

# enough about boring research, let's discuss trying generative AI

H.-T. Lin (NTU)

Two 2023 Attempts on Teaching in the Age of Generative AI.

# Attempt 1: Feeding Midterm Problems to ChatGPT

#### 2023.05.09 Data Structures and Algorithms Class

With the help of our TAs, we experimented using chatGPT to answer our midterm questions. I fed each problem (latex source) to chatGPT with the following prompt,

"Please solve the following problem within 200 words."

Then we asked the TAs to grade chatGPT's solutions using the same grading criteria as those your solutions received.



H.-T. Lin (NTU)

# **Details and Results**

- fixed, independent prompt for each problem
- no modifications, no repetitions
- 92 out of 200 points

-average of class: 125, human still wins? ;-)

- observations
  - chatGPT answered "popular problems" (e.g. those similar to Leetcode problems) pretty well.
  - chatGPT answered difficult "math problems" (e.g. big-O proof) pretty badly. (0 in those, actually)
  - chatGPT cannot easily deal with drawing problems (e.g. drawing a Tree). (at that time)
  - chatGPT dealt with "explanation problems" reasonably well.
     e.g. When the algorithm is fed with an array with a large length n that is already sorted, which algorithm runs the fastest? Explain your choice.

# Attempt 2: ChatGPT-Stimulated Discussions

#### 2023.09.06 Machine Learning Class

#### Experimental Items for HTML2023Fall

- at-home lecture watching
- physical discussions stimulated by chatGPT

ChatGPT 3.5 ~		<u>د</u>
5	You ML可以處理停機問题嗎	
1	ChatGPT 總器學習 (NUL) 本身不能將決停機問題 (haiting problem) 。 停機問題是計算單論中約一 個罪與問題,它試圖應定是否可以設計一個算法,該算法可以利用任何防定的算法是否會 在某個個人上你让快可愛無限證場,同常在否認 (Alonzo Church) 和印爾 祖童 (Alan Turing) 分別證明了停機問題的不可解性,即無法認計一個通過算法來將法這個問題。 總器學習是就容藝的時計方法,用於訓練模型以進行特定任務,何知圖會說時,語書謂 導,預理等,ML裡型位部於訓練發展,通過學習從整整中現現模式來執行任務,但是, ML模型仍然受到停機問題的限制,因為它們本算上是基於算法的,這直無法解決停機問 题。	
	停機問題是一個理論性的問題,與實際的機器學習應用無關。ML可以用於解決許多實際問題。但它無法解決停機問題,因為該問題已。此時是不可解的,要處理實際的問題,您可	
Mes	ssage ChatGPT 🕲 👩	
	ChatGPT can make mistakes. Consider checking important information.	(?)

#### chatGPT-stimulated discussions toward critical thinking

H.-T. Lin (NTU)

# **Details and Results**

- offline lecture watching, real-time lecture highlights, then do discussions
- selected questions, usually contextless ones, to ask chatGPT —maximize student experience, time control
- tried two modes:
  - chatGPT answers first (less pressure, sometimes prepared in advance)
  - I answer first (more exciting)

# Does It Help Learning?

- positive feedback for those who choose to come to classroom, but not for everyone
  - (student feedback) 我很喜歡老師這種課堂以補充和回答問題為主的 教學方式。原因是我認為課本上的知識都寫在課本上了,大家都可 以自己看,如果課堂沒有超出課本上的知識,那來上課的價值大大 降低。因此我認為老師對於課堂的設計非常合理。
  - (student feedback) 上課感覺好像是在快問快答,東西的質量沒有很高。我會比較喜歡安排過有系統的課程

# Summary with My Thoughts after Attempts

#### need research on process

#### certification challenge

trying is easy; systematic testing is difficult

-need research on evaluation

#### manipulation challenge

generating something is easy; generating good thing is difficult

-need research on control

let's research more to move GAI

to trustworthy commercial tools

# Appendix: Image Citations

Except for images from our papers and those cited on the fly, other images are cited as follows:

# Free Content Use from https:

//pixabay.com/vectors/robot-machine-technology-science-312566/

CC-0 from https://bioicons.com/icons/cc-0/Molecular\_modelling/ Simon\_Dürr/smiles.svg

#### CC-0 from

https://commons.wikimedia.org/wiki/File:Demonic\_possession.jpg



: Figure 1 from Song et al., ICLR 2021

https://openreview.net/forum?id=PxTIG12RRHS