Rethink Deep Learning with Invariance in Data Representation

A Tutorial at The Web Conference 2025 in Sydney (WWW 2025)

Shuren Qi[1], Fei Wang[2], Tieyong Zeng[1], and Fenglei Fan[3]

[1] CMAI, IMS, Department of Mathematics, The Chinese University of Hong Kong, HK
[2] Weill Cornell Medicine, Department of Population Health Sciences, Cornell University, USA
[3] Department of Data Science, City University of Hong Kong, HK

Homepage

Proposal, slides, reading list, video, and more materials available at

https://shurenqi.github.io/wwwtutorial/


Rethink Deep Learning with Invariance in Data Representation

Shuren Qi
The Chinese University of Hong Kong
Hong Kong, China

Fei Wang
Cornell University
New York, United States

Tieyong Zeng
The Chinese University of Hong Kong
Hong Kong, China

Fenglei Fan
The Chinese University of Hong Kong
Hong Kong, China

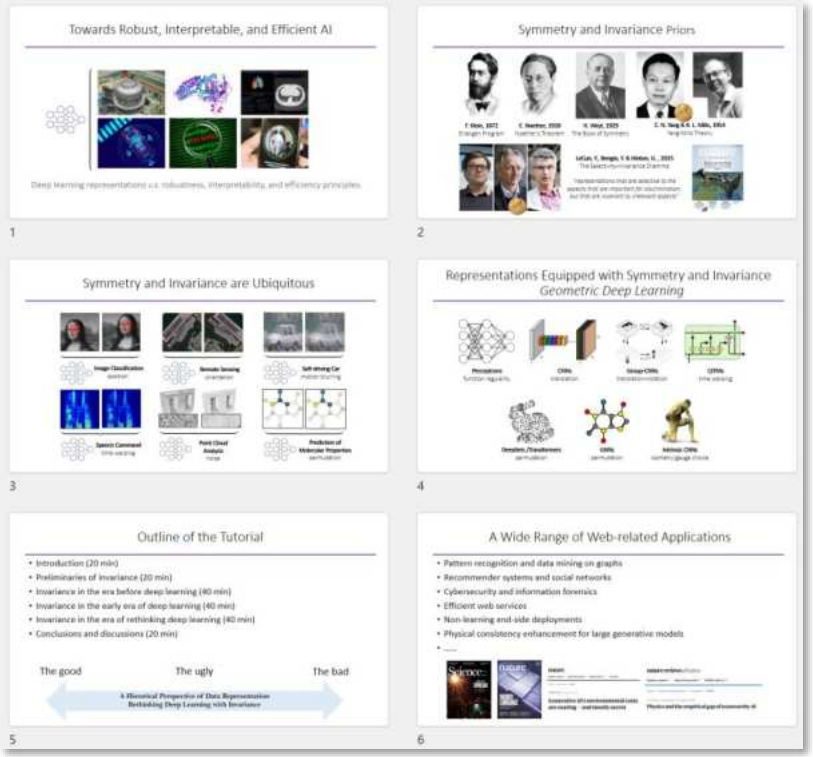
Abstract

CCS Concepts

Keywords

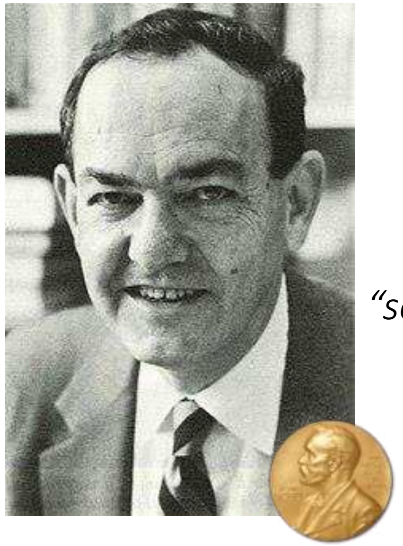
1 Topic and Relevance

2 Content



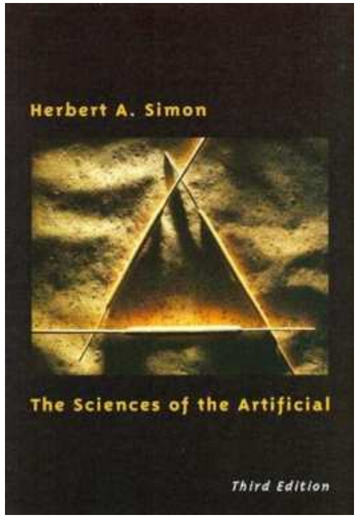
Deep (Representation) Learning,
A Big Bang Moment For AI

Data Representation



H. Simon, 1969
The Sciences of the Artificial

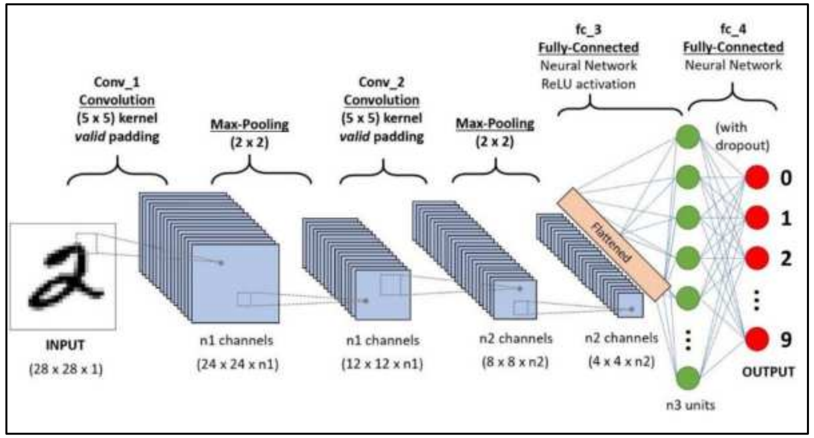
“solving a problem simply means representing it so as to make the solution transparent”

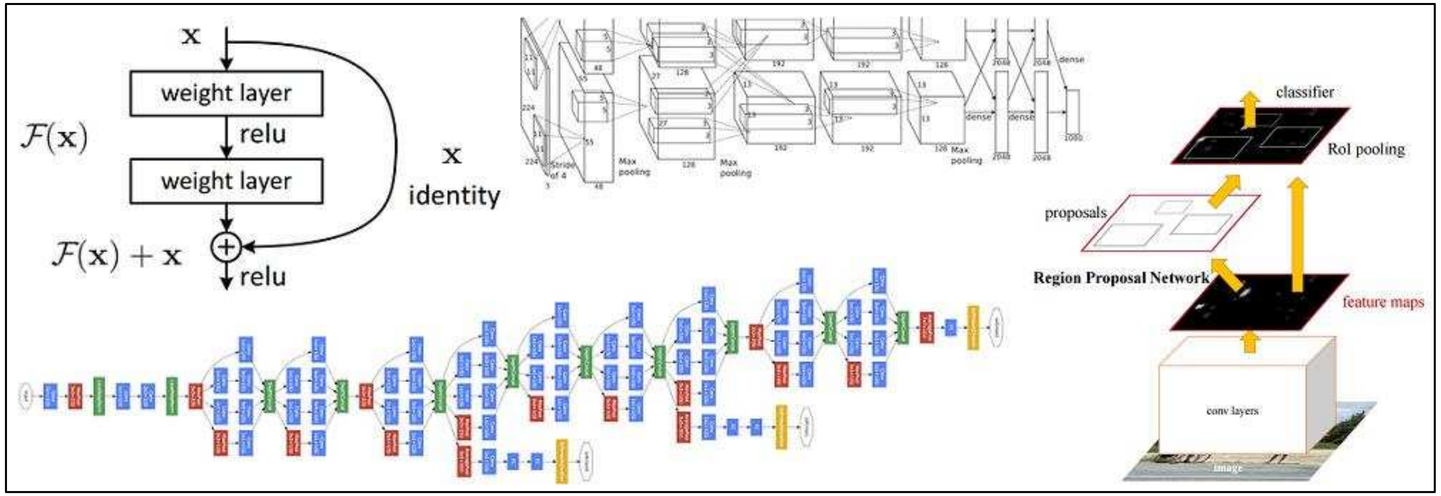


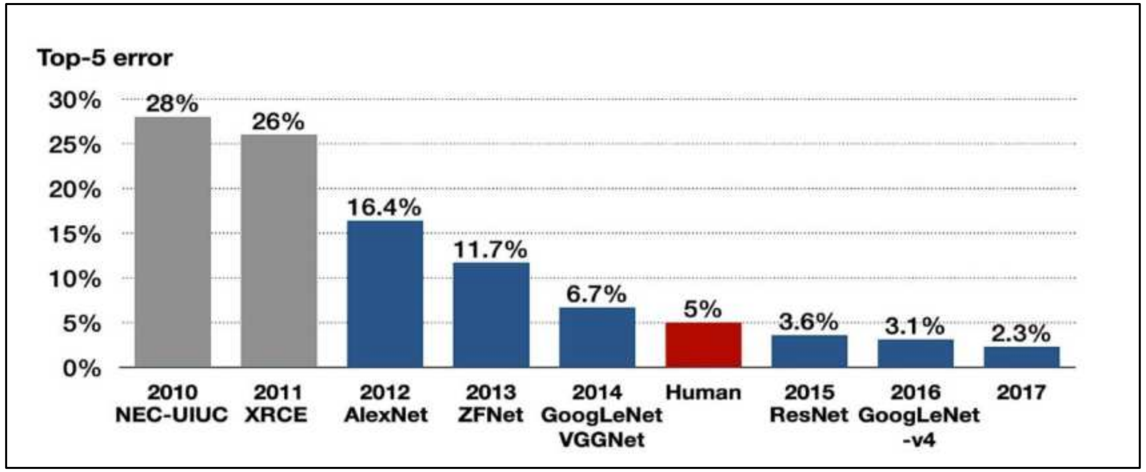
Data ⇨ Representation ⇨ Knowledge discovery ⇨ Application

- H Simon. *The Sciences of the Artificial (Third edition)*. MIT Press, 1996.

Processing Human Perceptual Information







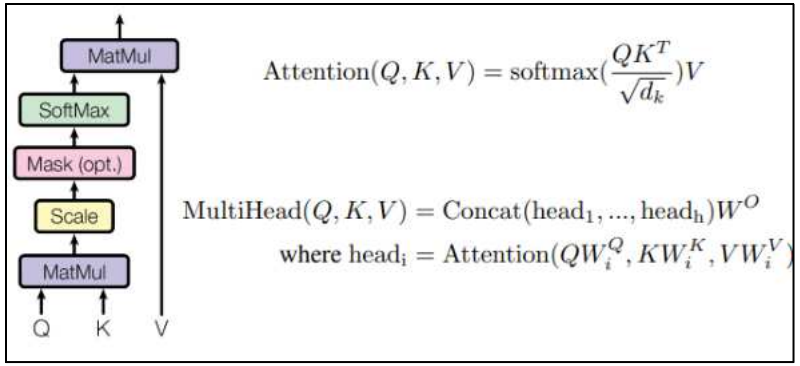
- J Deng, W Dong, R Socher, et al. ImageNet: A large-scale hierarchical image database. *CVPR,* 2009.

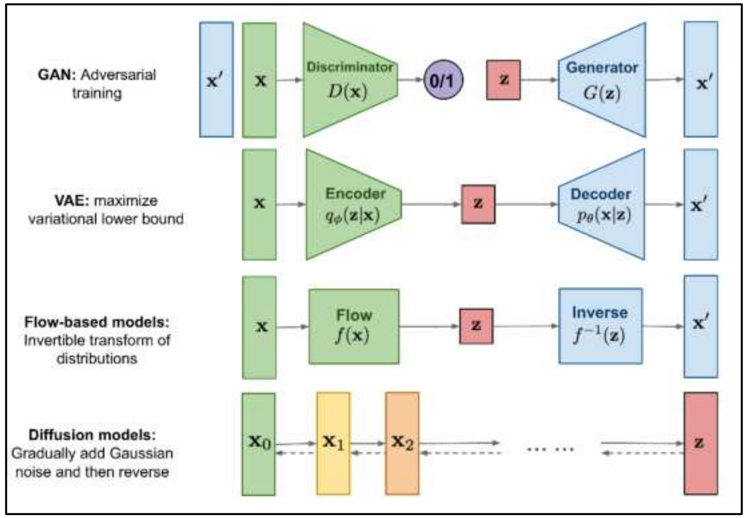
Generating Realistic Media

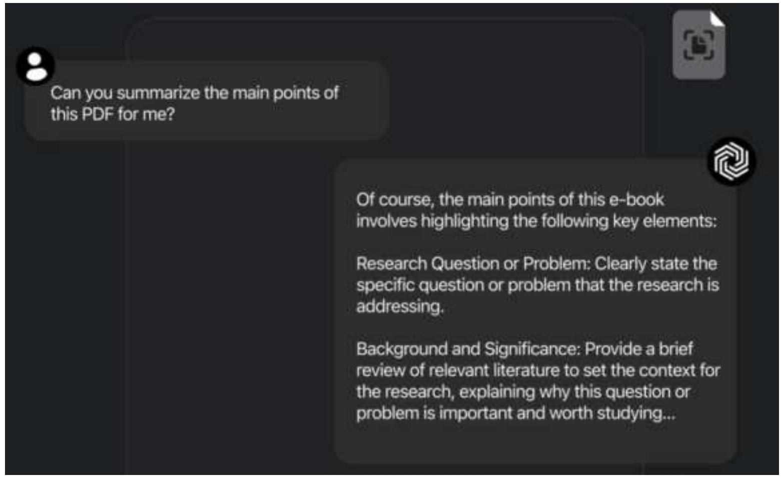
ChatGPT

DALL·E 3

OpenAI SORA



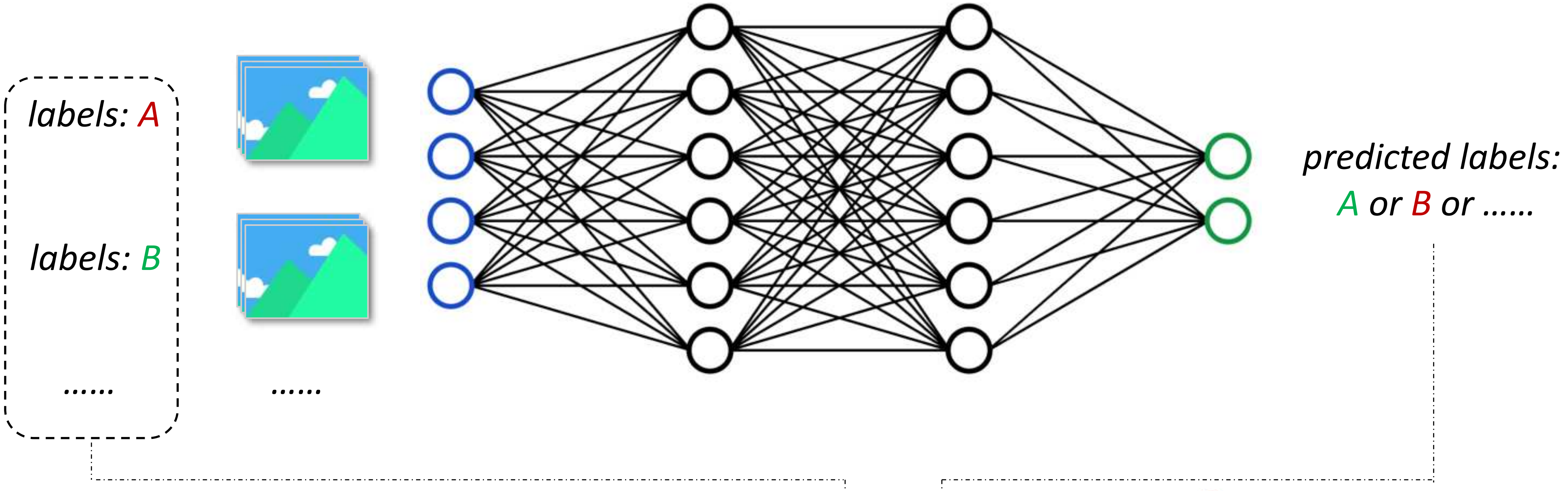




- Z Epstein, A Hertzmann, L. Herman, et al. Art and the science of generative AI. *Science*, 2023.

Empirical Risk Minimization (ERM), Behind All These Successes

Empirical Risk Minimization



$$\mathbb{E}_{\mathbf{x},\mathbf{y}\in S}[\ell(\mathbf{y}, h(\mathbf{x}))]$$

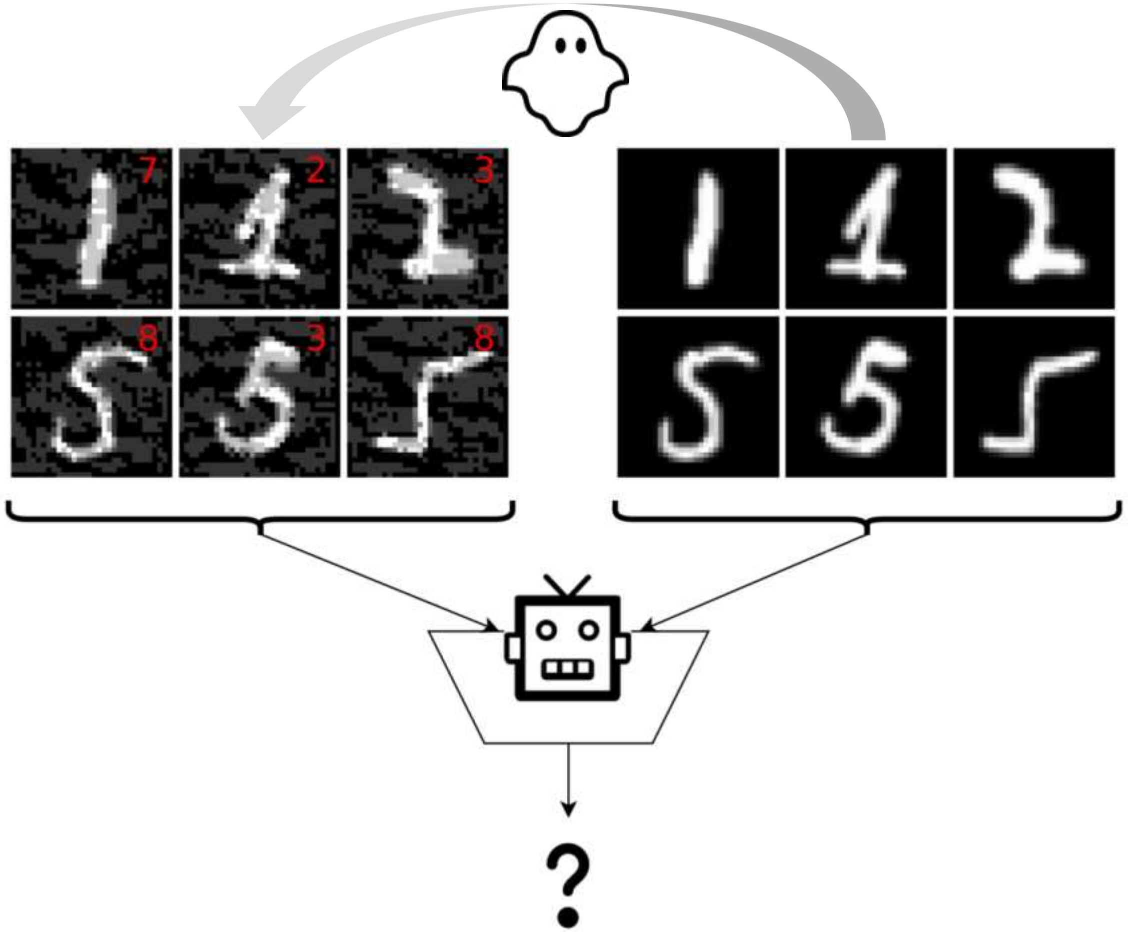
Empirical learning draws the lines between categories.

But what about robustness, interpretability, and efficiency?

- V Vapnik. Principles of risk minimization for learning theory. *NIPS,* 1991.

Robustness of Empirical Learning

- Robustness: the performance of a system is stable for intra-class variations on the input.



Universal

Color

One-pixel

Watermark

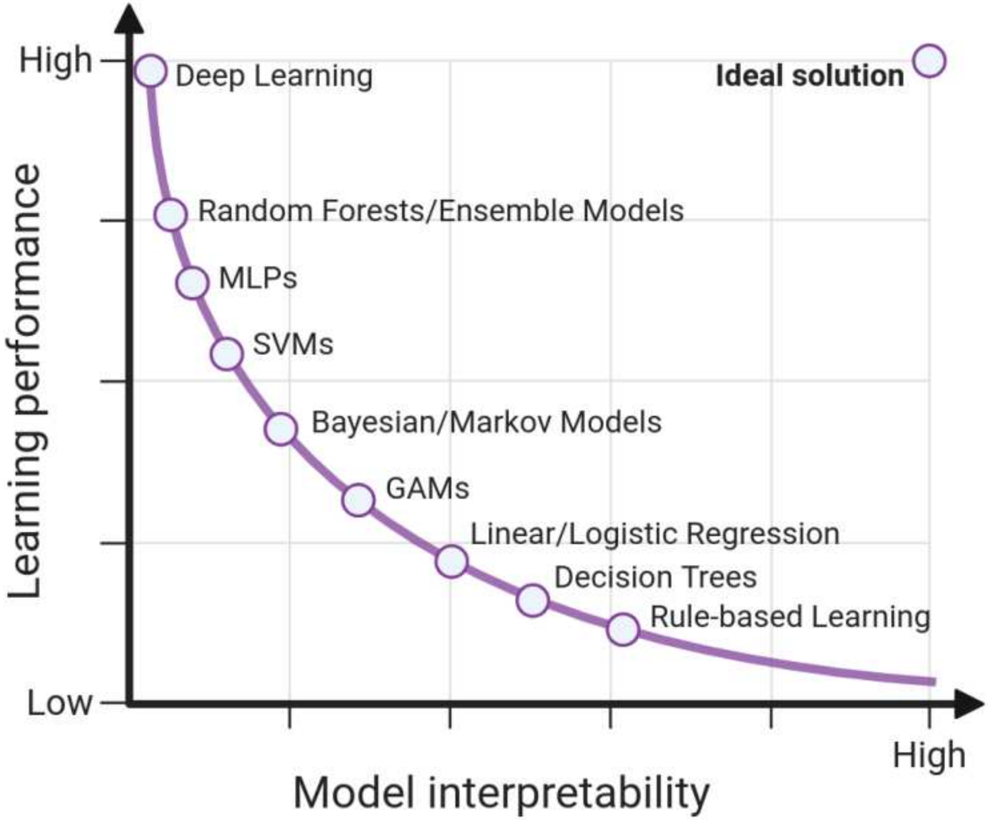
Physical

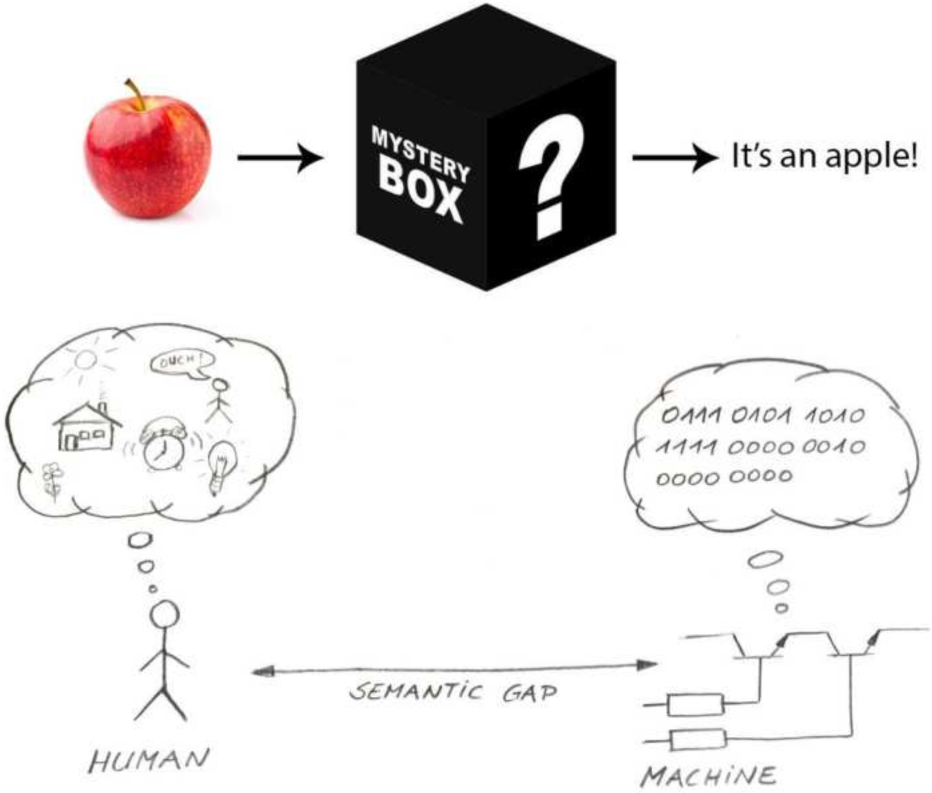
Weird

- C Buckner. Understanding adversarial examples requires a theory of artefacts for deep learning. *Nature Machine Intelligence*, 2020.

Interpretability of Empirical Learning

- Interpretability: the behavior of a system can be understood or predicted by humans.

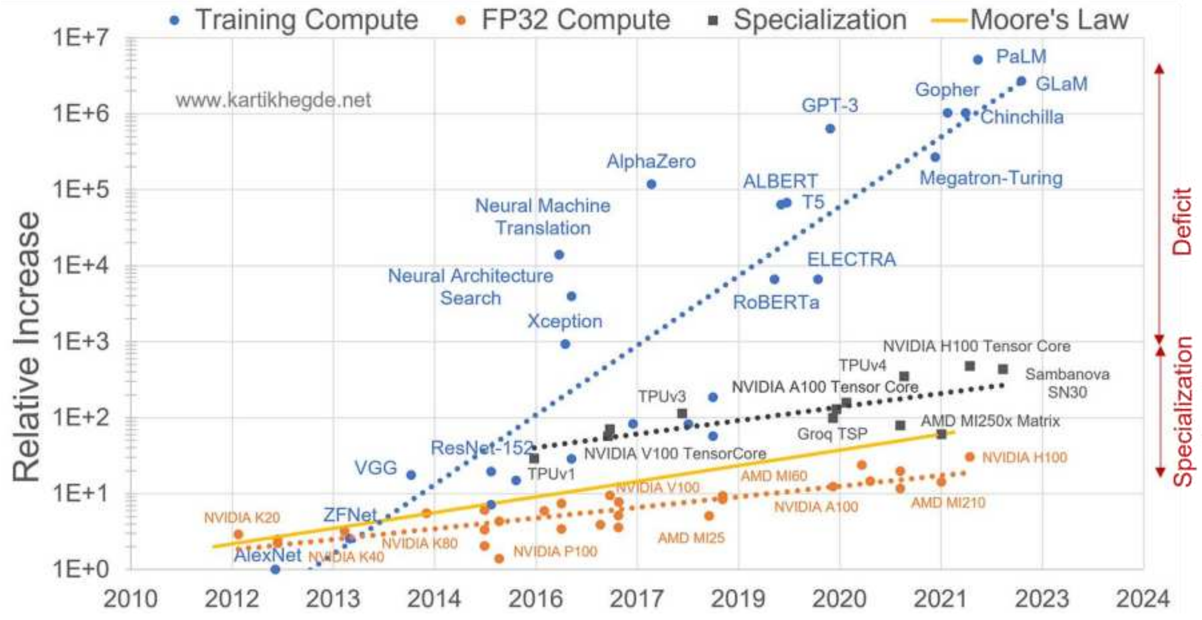


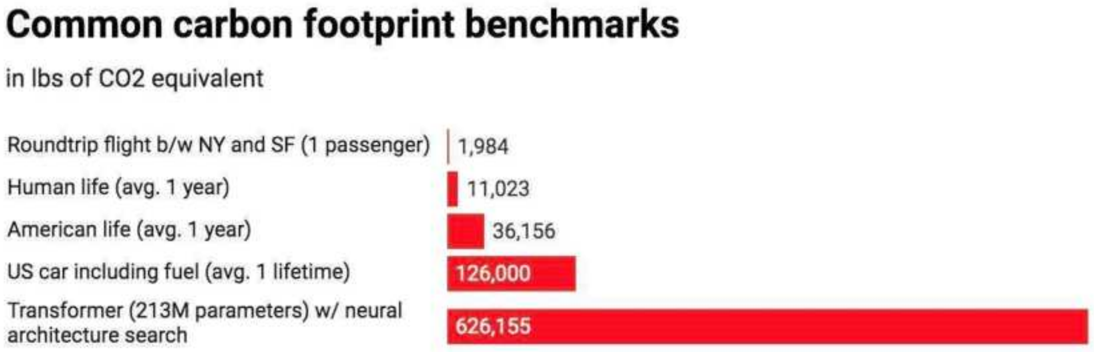


- X Li, C Cao, Y Shi, et al. A survey of data-driven and knowledge-aware explainable AI. *TKDE,* 2020.

Efficiency of Empirical Learning

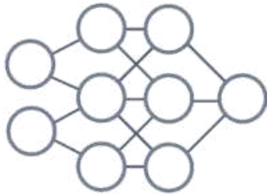
- Efficiency: the real-time availability and energy cost during human-computer interaction.

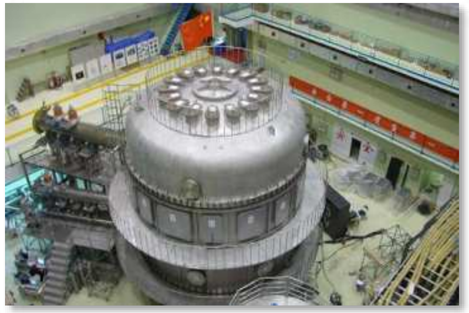




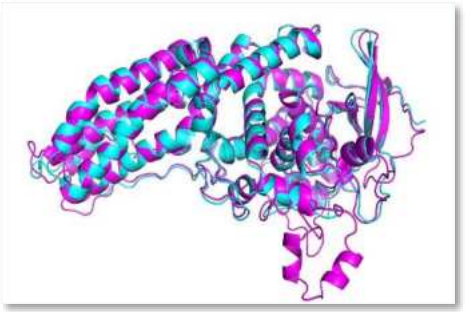
- E Strubell, A Ganesh, A McCallum, et al. Energy and policy considerations for modern deep learning research. *AAAI,* 2020.

When Moving Towards Trustworthy AI





Tokamak Control



Drug Discovery

Imaging Diagnostics

Automatic Driving

Cyber Security

Biometrics

Empirical learning *v.s.* robustness, interpretability, efficiency…

- H Liu, M Chaudhary, H Wang. Towards trustworthy and aligned machine learning: A data-centric survey with causality perspectives. *arXiv preprint arXiv:2307.16851*, 2023.

A Foundational Prior Underlying Both Natural World And AI Systems

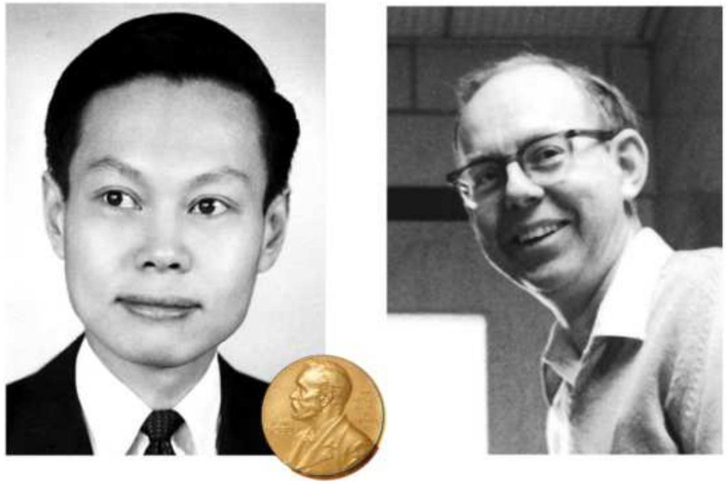
Invariance/Symmetry in Natural World

- A symmetry of a system is a transformation that leaves a certain property invariant.

F. Klein, 1872
Erlangen Program

E. Noether, 1918
Noether's Theorem

H. Weyl, 1929
The Book of Symmetry



C. N. Yang & R. L. Mills, 1954
Yang-Mills Theory

- F Klein. A comparative review of recent researches in geometry. *Bulletin of the American Mathematical Society*, 1893.
- H Weyl. *Symmetry.* Princeton University Press, 2015.

Invariance/Symmetry in AI Systems

- An AI system is a digital modeling of the physical systems in the natural world.

Y. LeCun, Y. Bengio & G. Hinton, 2015, Deep learning, Nature

The Selectivity–Invariance Dilemma:
“representations that are selective to the aspects that are important for discrimination, but that are invariant to irrelevant aspects”

- Y LeCun, Y Bengio, G Hinton. Deep learning. *Nature*, 2015.
- Y Bengio, A Courville, P Vincent. Representation learning: A review and new perspectives. *TPAMI*, 2013.

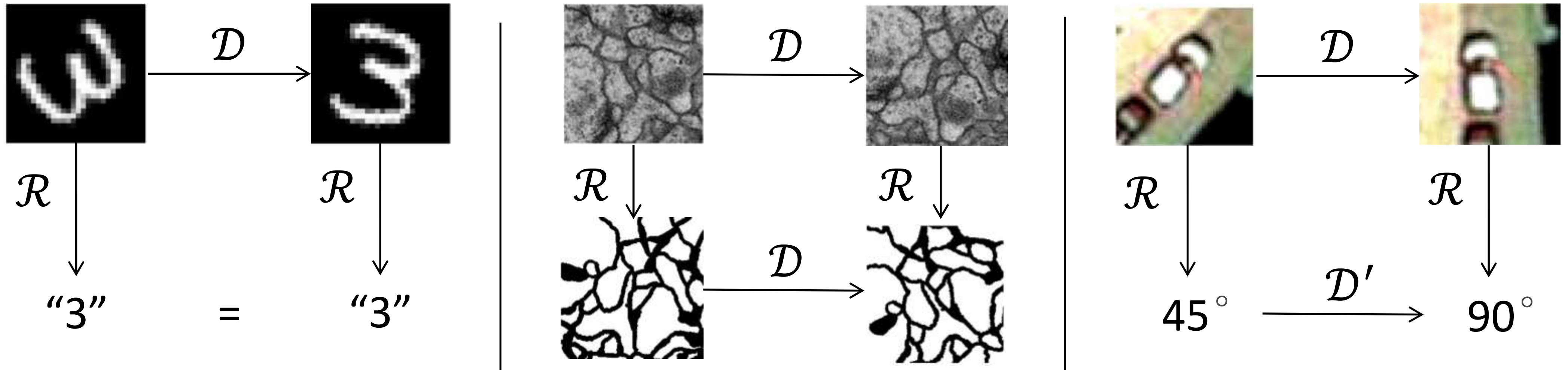
How Invariance/Symmetry Helps Robustness, Interpretability, Efficiency

- Perfect robustness — the performance of the AI system remains invariant with respect to the transformations of interest.
- Interpretable concept — humans and AI systems share a basic concept that allows humans to predict AI behavior on transformations of interest.
- Structural efficiency – AI systems no longer need to memorize non-discriminative data variants.

A Formalization of Invariance/Symmetry (in Representation)

- Invariance: $\mathcal{R}(\mathcal{D}(f)) \equiv \mathcal{R}(f)$
- Equivariance: $\mathcal{R}(\mathcal{D}(f)) \equiv \mathcal{D}(\mathcal{R}(f))$
- Covariance: $\mathcal{R}(\mathcal{D}(f)) \equiv \mathcal{D}'(\mathcal{R}(f))$
- $\mathcal{R}$ is a representation, $\mathcal{D}$ is a degradation, and invariance/equivariance is a special case of covariance with $\mathcal{D}' = \mathrm{id}/\mathcal{D}$

- K Lenc, A Vedaldi. Understanding image representations by measuring their equivariance and equivalence. *CVPR*, 2015.



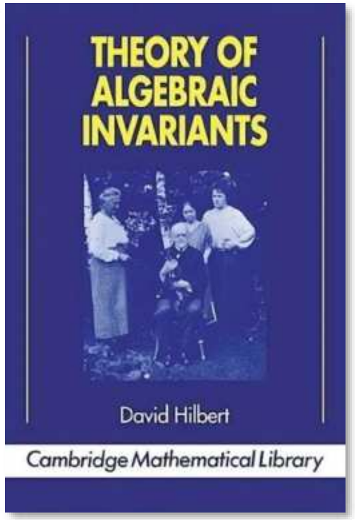
Invariance
$\mathcal{D}$ = rotate, $\mathcal{R}$ = classifier

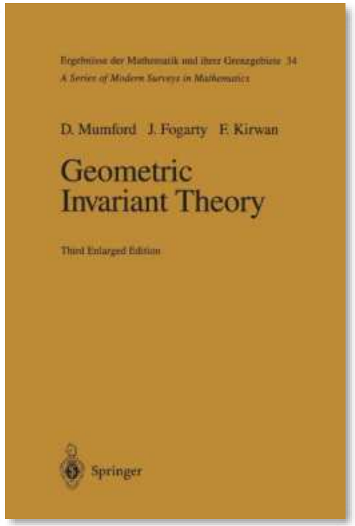
Equivariance
$\mathcal{D}$ = rotate, $\mathcal{R}$ = detector

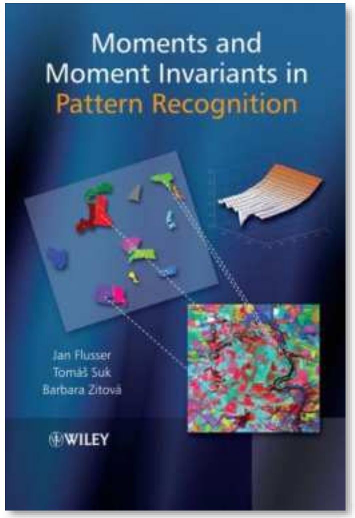
Covariance
$\mathcal{D}$ = rotate, $\mathcal{R}$ = estimator

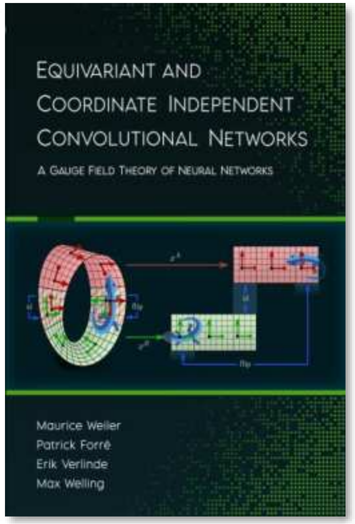
A History of Invariance/Symmetry (in Representation)

Algebraic Invariants	Geometric Invariants	Moment Invariants	Multiscale and Wavelet	CNN to Geometry Deep Learning
1840s	*1960s*	*2000s*	*2010s*	*2020s*
Hilbert Cayley Klein…	Mumford	Flusser	Lowe Lindeberg Mallat…	LeCun Bronstein…









What I Did With My Collaborators In The Process Of Invariance?

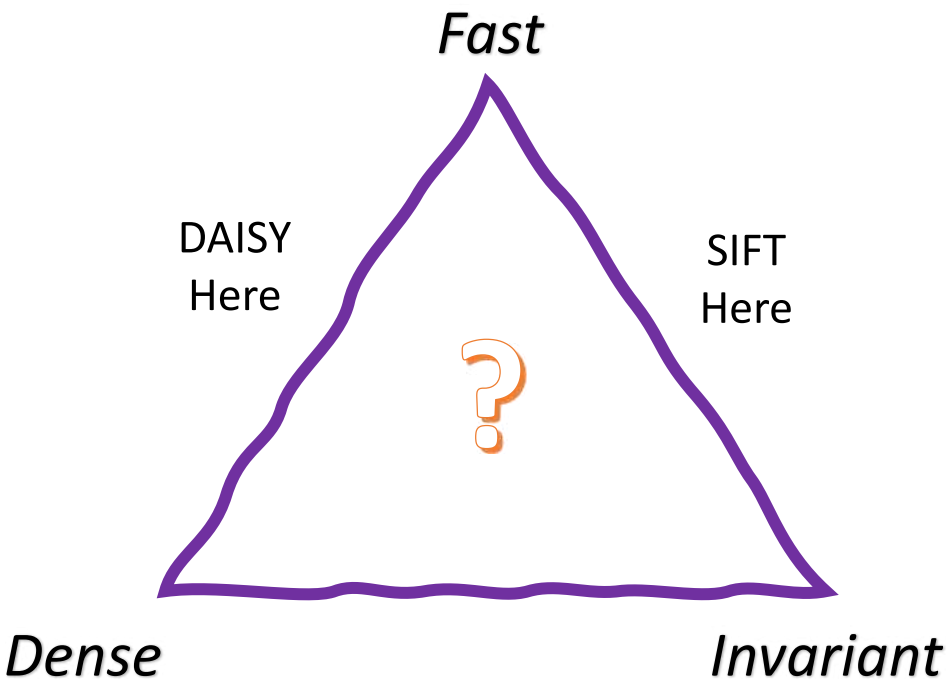
Our Contributions

	Moment Invariants	Multiscale and Wavelet	CNN to Geometry Deep Learning
	2000s	*2010s*	*2020s*
	Flusser	Lowe Lindeberg Mallat…	LeCun Bronstein…
Assumption	Global	Local	Hierarchical
Invari./Discri.	Invariance	Mid?	Discriminability
Robust./Interp./Effic.	Good	Mid?	Not Good
Our Contributions	CSUR'23 TPAMI'24 TIP review	TPAMI'23 USENIX Security'25	ICCV'25 TIP major

Designing Local Invariants

- Reviewing the local invariants, we note a gap: no method is fast, invariant, and dense.
- We tried to define truly dense invariants while being fast enough. We achieved this goal by exploring the potential of **classical moment invariants**.


A Principled Design of Image Representation: Towards Forensic Tasks

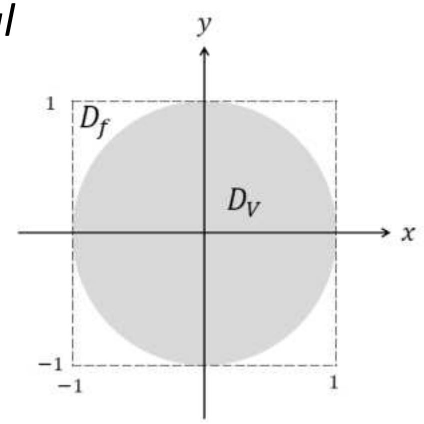


- S. Qi, Y. Zhang, C. Wang, et al. A Principled Design of Image Representation: Towards Forensic Tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2023, 45(5): 5337 - 5354

Moments: From Global to Local

- First, we extend the definition of classical moments from the global to the local with scale space. Here, local coordinate system (x', y') is a translated and scaled version of the global coordinate system (x, y), with translation offset (u, v) and scale factor w.
- Two interesting properties: generic nature and local representation capability.

Global



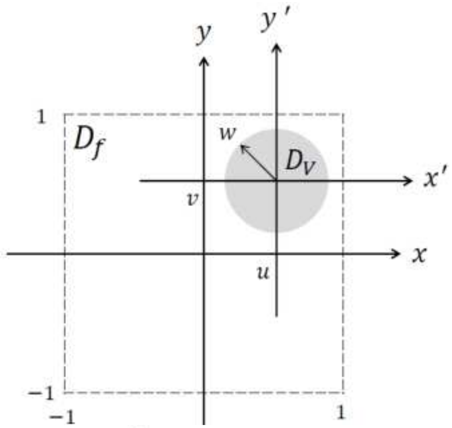
$$\langle f, V_{nm} \rangle = \iint_D R_n^*(r) A_m^*(\theta) f(r, \theta) r dr d\theta$$

Our Transformations

$$(x', y') = \frac{(x, y) - (u, v)}{w}$$

$$\begin{cases} r' = \sqrt{(x')^2 + (y')^2} = \frac{1}{w}\sqrt{(x-u)^2 + (y-v)^2} \\ \theta' = \arctan(\frac{y'}{x'}) = \arctan(\frac{y-v}{x-u}) \end{cases}$$

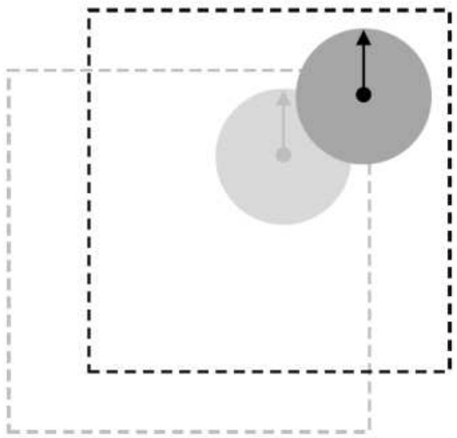
Local



$$\langle f, V_{nm}^{uvw} \rangle = \iint_D \underbrace{R_n^*(\overbrace{\frac{\sqrt{(x-u)^2 + (y-v)^2}}{w}}^{r'}) A_m^*(\overbrace{\arctan(\frac{y-u}{x-v})}^{\theta'})}_{(V_{nm}^{uvw}(x,y))^*} f(x, y) dx dy$$

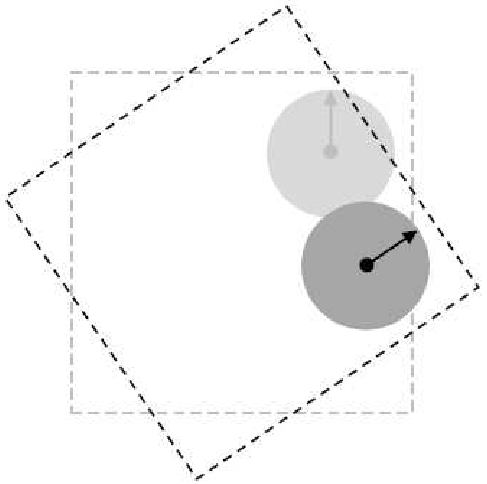
Moment Invariants: From Global to Local

- Then, we found the symmetry properties of the local definition for several geometric transformations.
- Therefore, rotation and flipping invariants can be obtained by taking the absolute values; translation and scaling invariants can be obtained by pooling over the $(u, v)/w$.



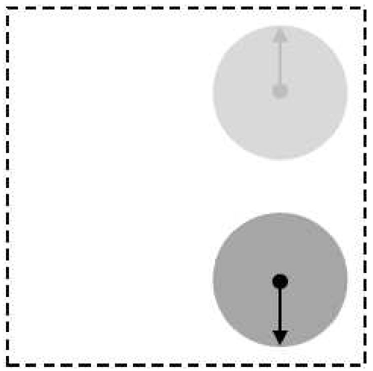
$$\langle f(x+\Delta x, y+\Delta y), V_{nm}^{uvw}(x,y)\rangle = \left\langle f(x,y), V_{nm}^{(u+\Delta x)(v+\Delta y)w}(x,y)\right\rangle$$

Translation Equivariance
w.r.t. (u, v)



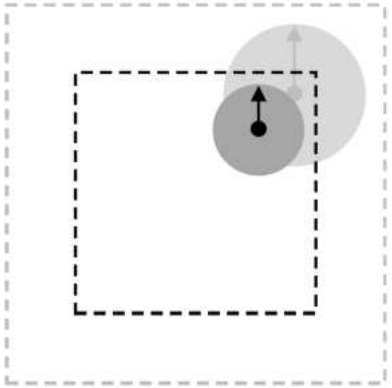
$$\langle f(r,\theta+\phi), V_{nm}^{uvw}(r',\theta')\rangle = \langle f(r,\theta), V_{nm}^{uvw}(r',\theta')\rangle\, A_m^*(-\phi)$$

Rotation Invariance
w.r.t. absolute values



$$\langle f(r,-\theta), V_{nm}^{uvw}(r',\theta')\rangle = (\langle f(r,\theta), V_{nm}^{uvw}(r',\theta')\rangle)^*$$

Flipping Invariance
w.r.t. absolute values

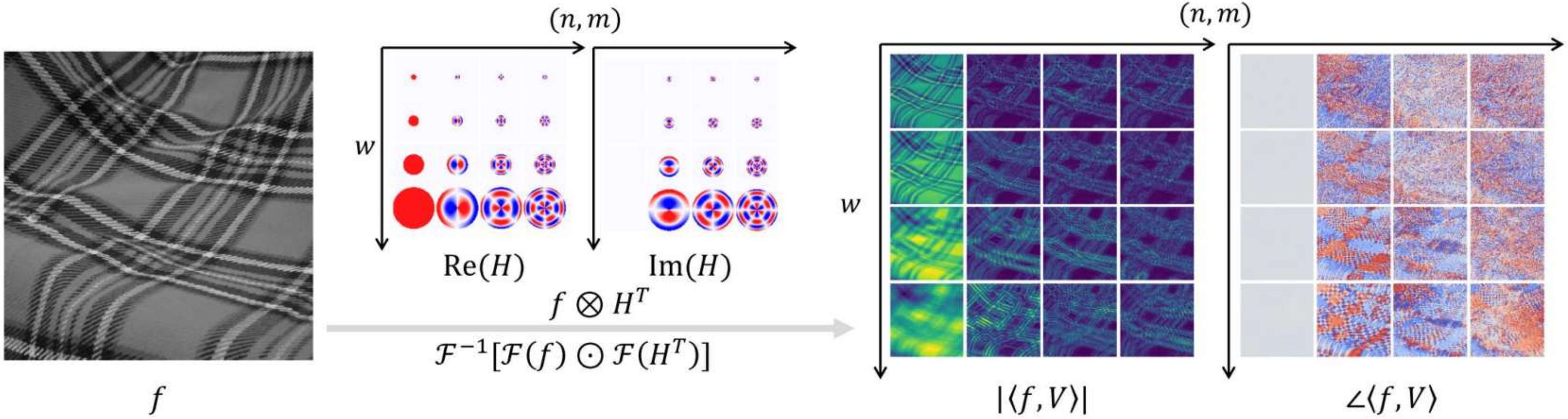


$$\langle f(sx,sy), V_{nm}^{uvw}(x,y)\rangle = \left\langle f(x,y), V_{nm}^{uv(ws)}(x,y)\right\rangle$$

Scaling Covariance
w.r.t. w

Fast Implementation

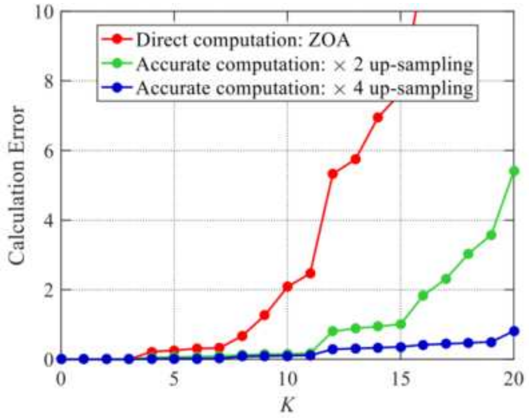
- Finally, we give a fast implementation by the convolution theorem.

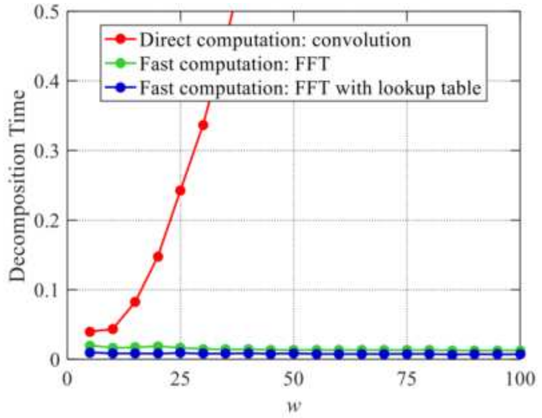


$\mathcal{O}({w_{\max}}^2 \#_{uv} \#_w)$ VS $\mathcal{O}(\#_w \#_{uv} \log \#_{uv})$

⇩

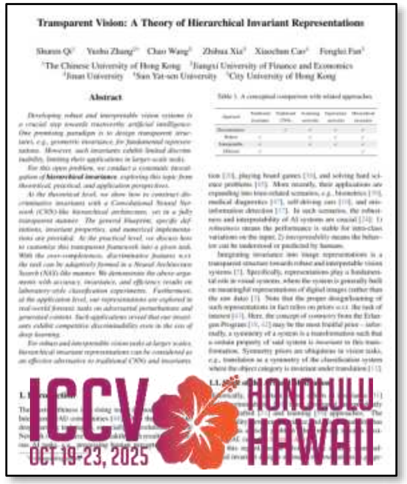
${w_{\max}}^2$ VS $\log(\#_{uv})$



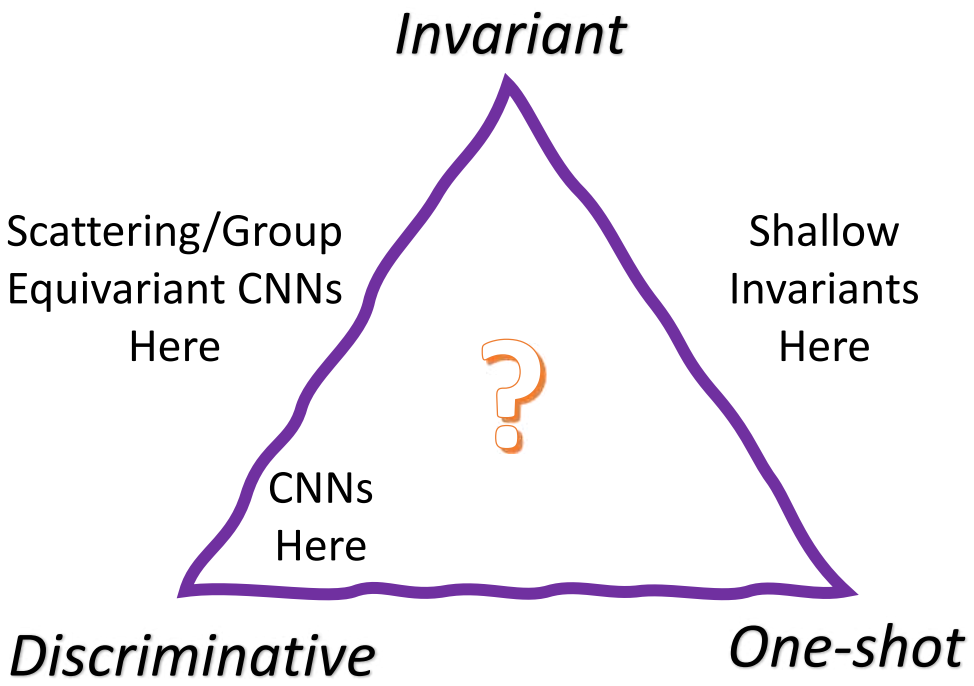


Exploring Hierarchical Invariants

- Reviewing the hierarchical invariants, we note a gap: no method is invariant, discriminative, and one-shot.
- We tried to define hierarchical (discriminative) invariants while being one-shot. We achieved this goal by exploring the potential of **classical moment invariants**.

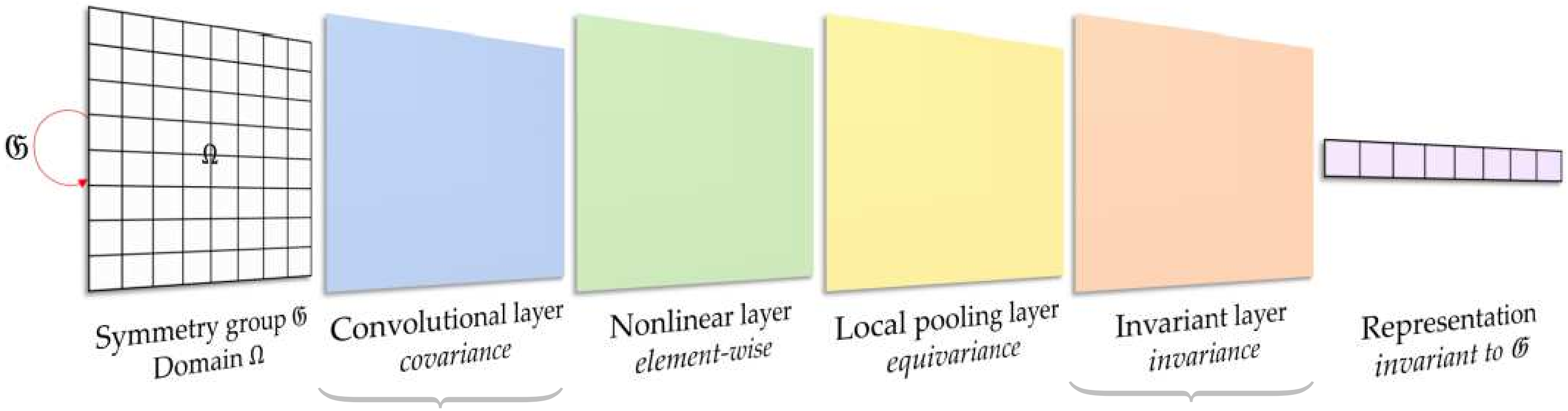


- S. Qi, Y. Zhang, C. Wang, et al. Transparent Vision: A Theory of Hierarchical Invariant Representations. *ICCV*, 2025.



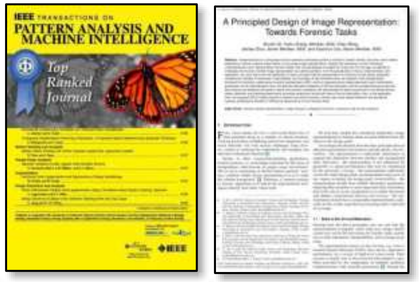
Blueprint

- First, we rethink the typical modules of CNN, unifying the theory of global and local invariants into a hierarchical network.

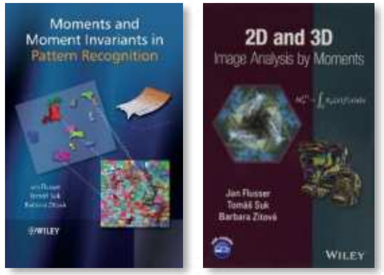


Need Local Invariants

Need Global Invariants



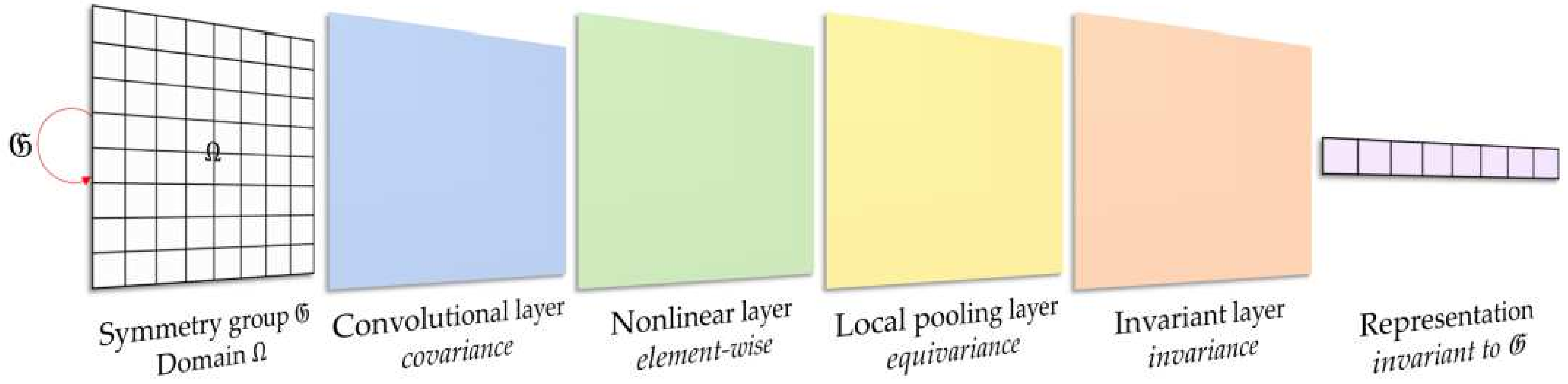
Recalling the geometric deep learning blueprint, we are surprised that we already have the components to form the hierarchical invariance, we just have not yet assembled them.



- S. Qi, Y. Zhang, C. Wang, et al. A Principled Design of Image Representation: Towards Forensic Tasks. *TPAMI*, 2023.
- J. Flusser, B. Zitova, T. Suk. *Moments and Moment Invariants in Pattern Recognition*. John Wiley & Sons, 2009.

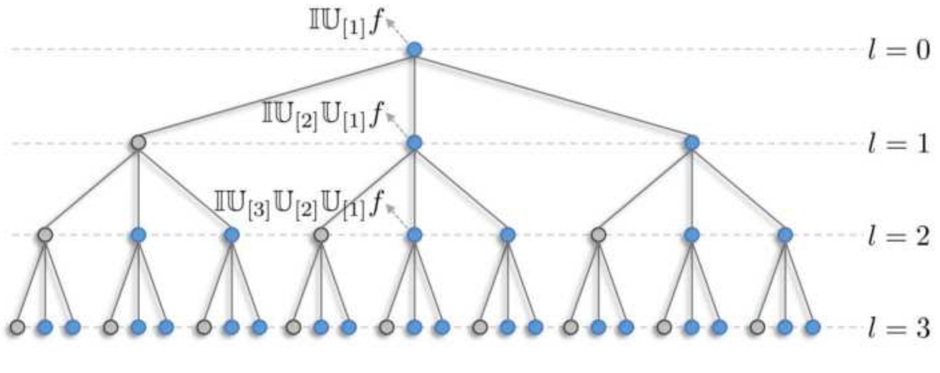
Definition

- Then, we can define new modules with their cascades to fulfill the blueprint:
  - Ω is 2D grid for images; $\mathfrak{G}$ is a translation, rotation, flipping, and scaling symmetry group over Ω.
  - **$\mathfrak{G}$-covariant convolutional layer:** $\mathbb{C}M \triangleq \langle M, V_{nm}^{uvw} \rangle = M(i,j;k) \otimes (H_{nm}^{w}(i,j))^T$
  - **Nonlinearity layer:** $\mathbb{S}M = \sigma(M(i,j)) \triangleq |M(i,j;k)|$
  - **Local pooling layer:** $\mathbb{P}M = M'$
  - **$\mathfrak{G}$-invariant layer:** $\mathbb{I}M = \mathcal{I}(\{\langle M(i,j;k), V_{nm}(x_i, y_j) \rangle\})$
  - **$\mathfrak{G}$-invariant representation:** $\mathcal{R}_p \triangleq \mathbb{I} \circ \mathbb{P}_{[L]} \circ \mathbb{S}_{[L]} \circ \mathbb{C}_{[L]} \circ \cdots \circ \mathbb{P}_{[1]} \circ \mathbb{S}_{[1]} \circ \mathbb{C}_{[1]}$

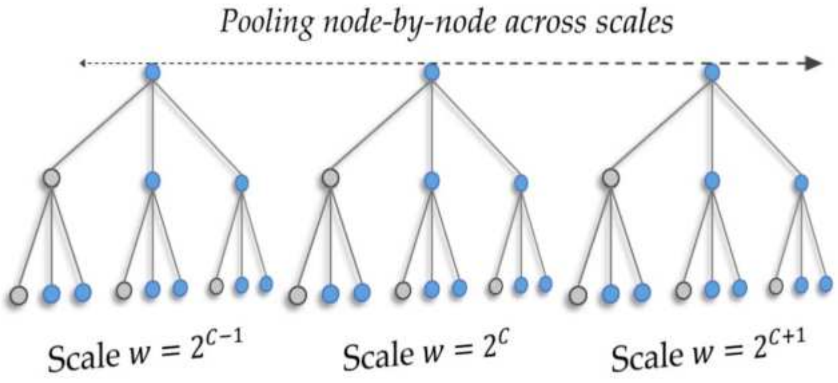


Property

- The group theory shows the one-shot symmetry property at each inter layer:
  - $\mathfrak{G}_1$ is the translation, rotation, and flipping symmetry group; $\mathfrak{G}_2$ is a scaling symmetry group, with scaling factor s. Any $\mathfrak{G}_0 \subseteq \mathfrak{G}_1 \times \mathfrak{G}_2$ as the symmetry group of interest. A representation unit denoted as $\mathbb{U} \triangleq \mathbb{P} \circ \mathbb{S} \circ \mathbb{C}$.
  - $\mathfrak{G}_1$ **Equivariance:** $\mathbb{U}_{[L]} \circ \cdots \circ \mathbb{U}_{[2]} \circ \mathbb{U}_{[1]}(\mathfrak{g}_1 M) \equiv \mathfrak{g}_1 \mathbb{U}_{[L]} \circ \cdots \circ \mathbb{U}_{[2]} \circ \mathbb{U}_{[1]}(M)$
  - $\mathfrak{G}_2$ **Covariance:** $\mathbb{U}^w_{[L]} \circ \cdots \circ \mathbb{U}^w_{[2]} \circ \mathbb{U}^w_{[1]}(\mathfrak{g}_2 M) \equiv \mathfrak{g}'_2 \mathbb{U}^w_{[L]} \circ \cdots \circ \mathbb{U}^w_{[2]} \circ \mathbb{U}^w_{[1]}(M) \quad \mathfrak{g}'_2 \mathbb{U}^w \triangleq \mathfrak{g}_2 \mathbb{U}^{ws}$
  - $\mathfrak{G}_0$ **Hierarchical Invariance:** $\mathbb{I}(\mathfrak{g}'_0 M)_{[L]} \equiv \mathbb{I} M_{[L]}$

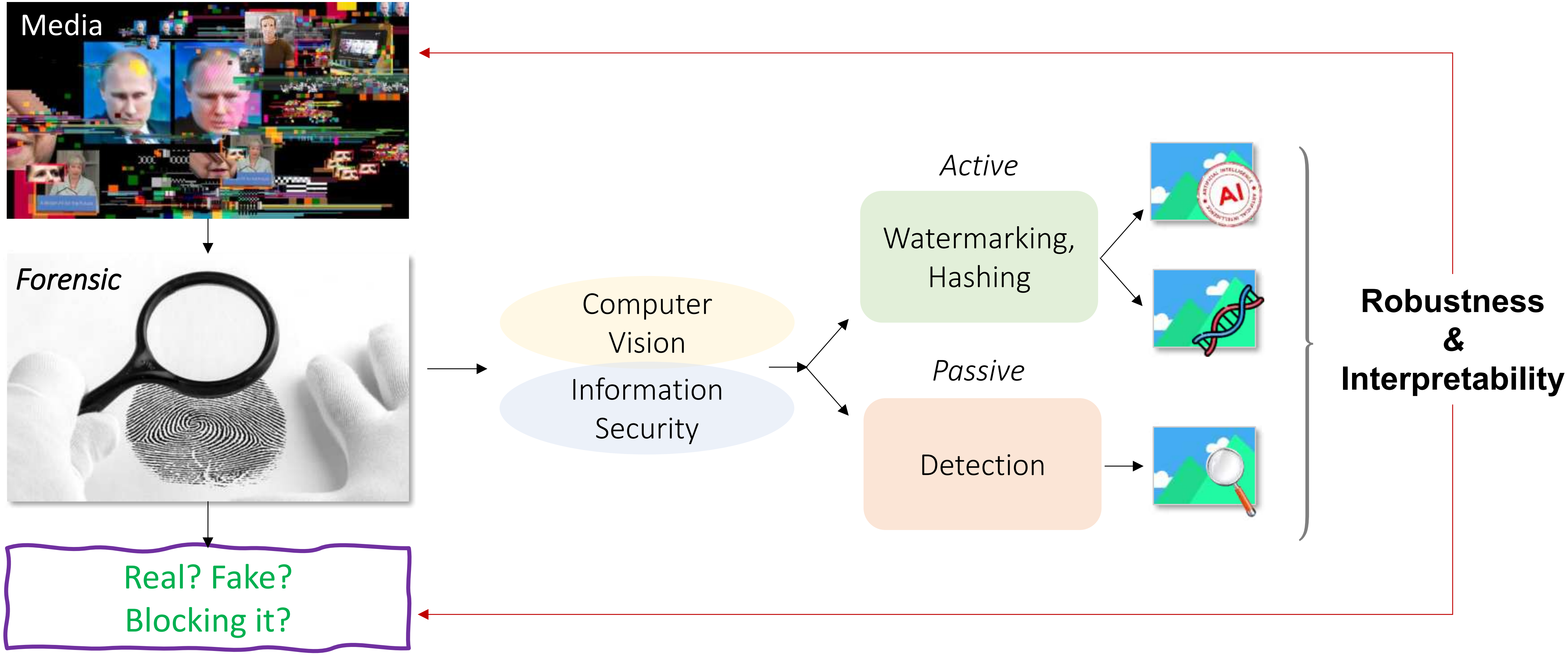


A Single-scale Practice with $\mathfrak{G}_1$ Invariance



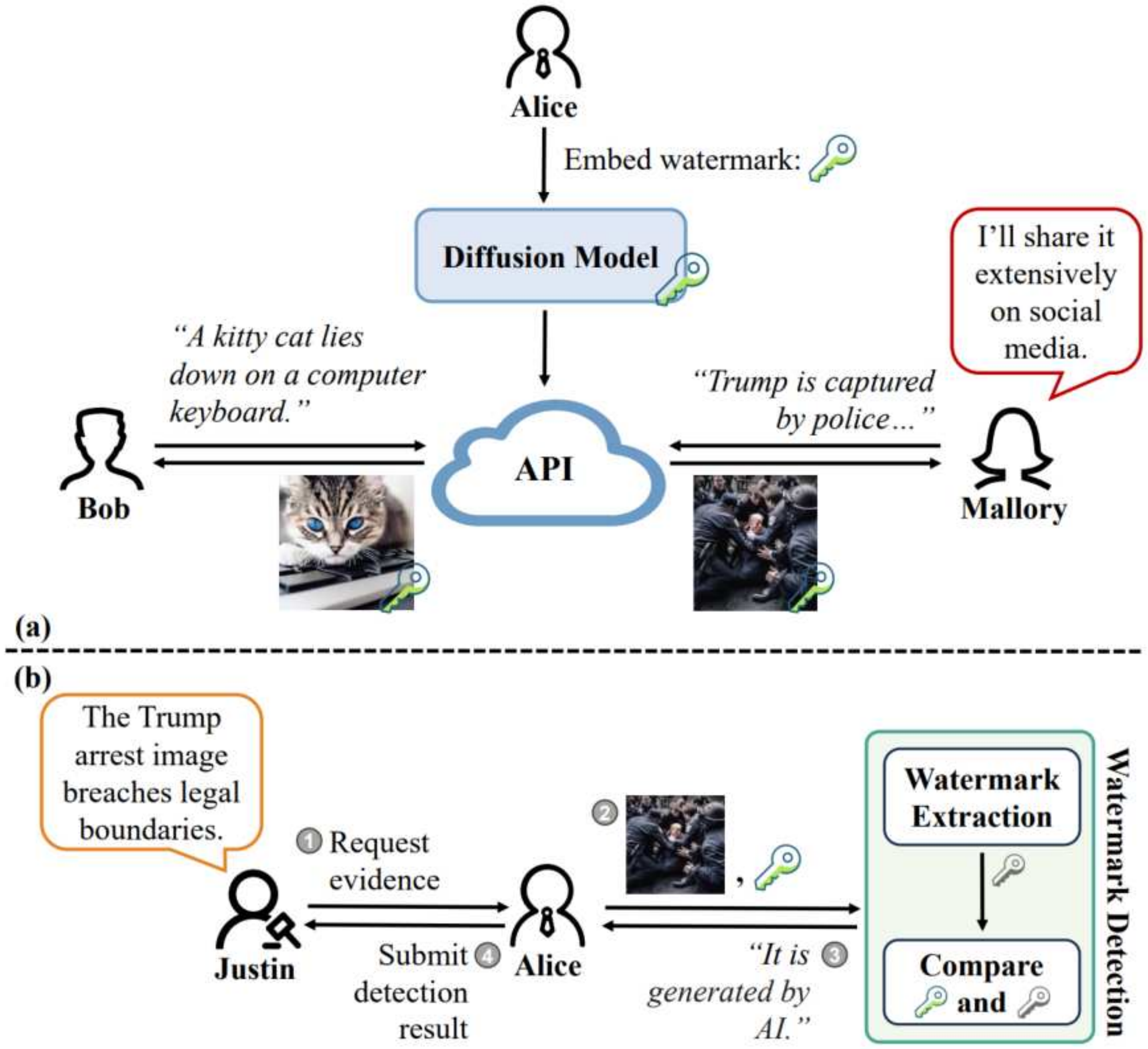
A Multi-scale Practice with $\mathfrak{G}_0$ Invariance

Forensic, Fighting Against AIGC Abuse



- T Wang, Y Zhang, S Qi, et al. Security and privacy on generative data in AIGC: A survey. *ACM Computing Surveys*, 2024.

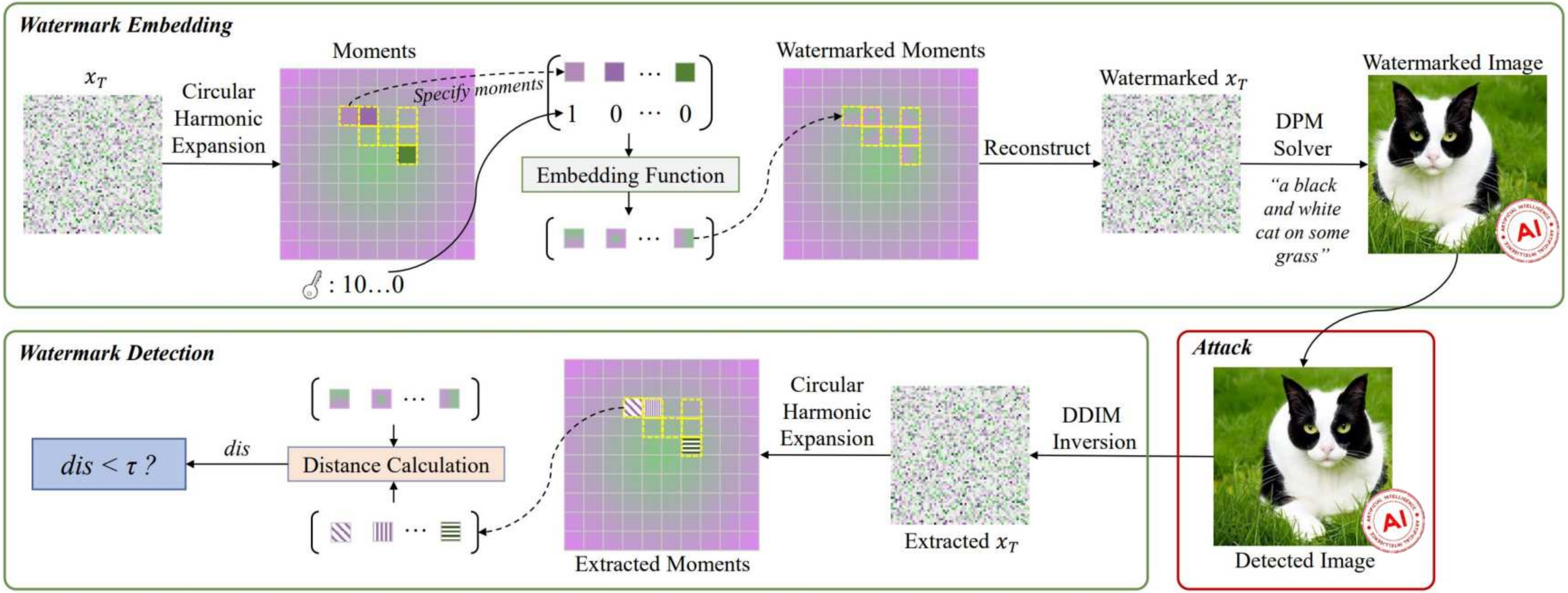
AIGC Watermarking: Motivations



Is there a balance between robustness and imperceptibility?

- Y. Zhang, M. Shen, S. Qi*, et al. Kill Two Birds with One Stone: Balance Imperceptibility and Robustness in Diffusion Model Watermarking. *TIP review*, 2025.

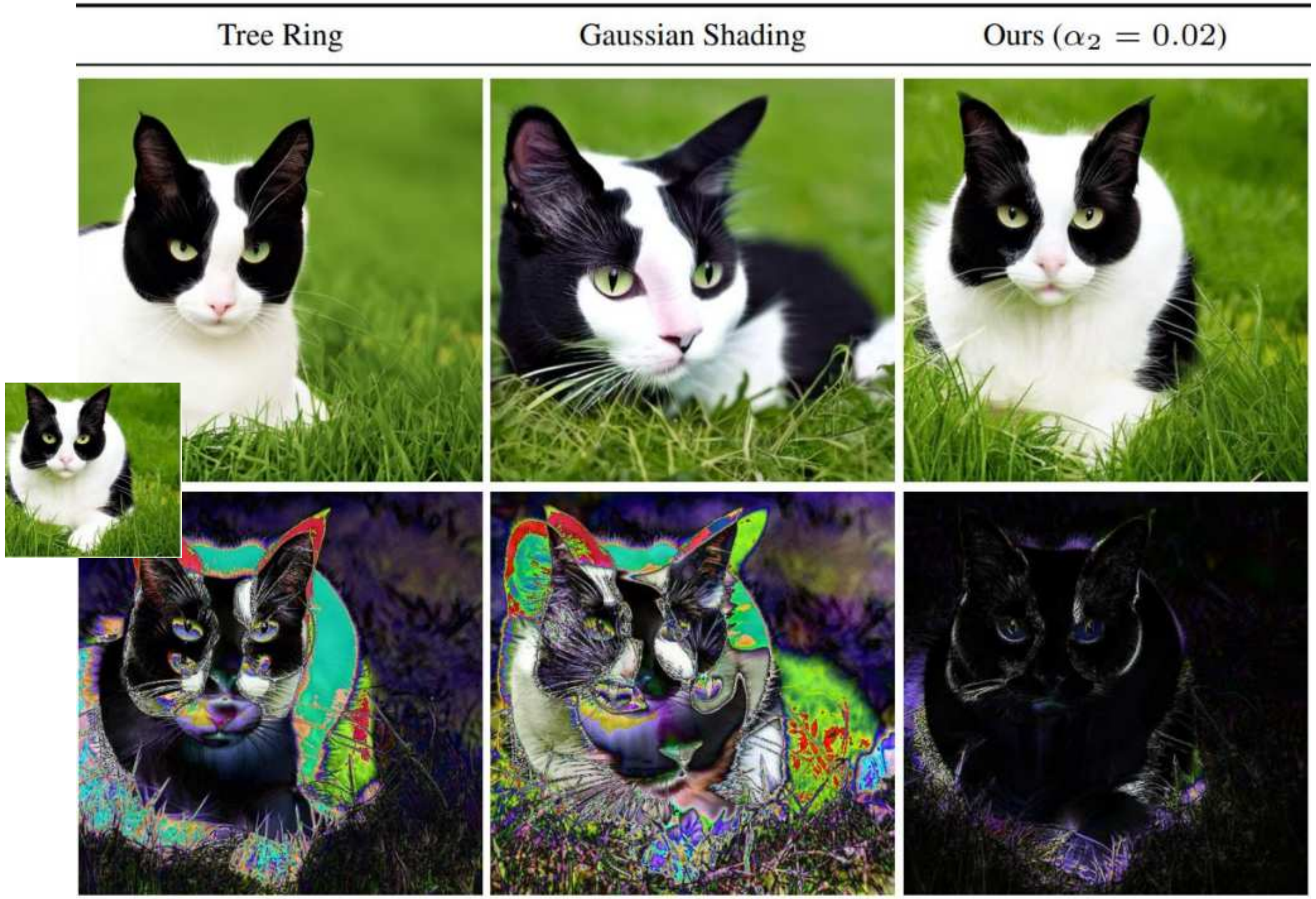
AIGC Watermarking: Ideas



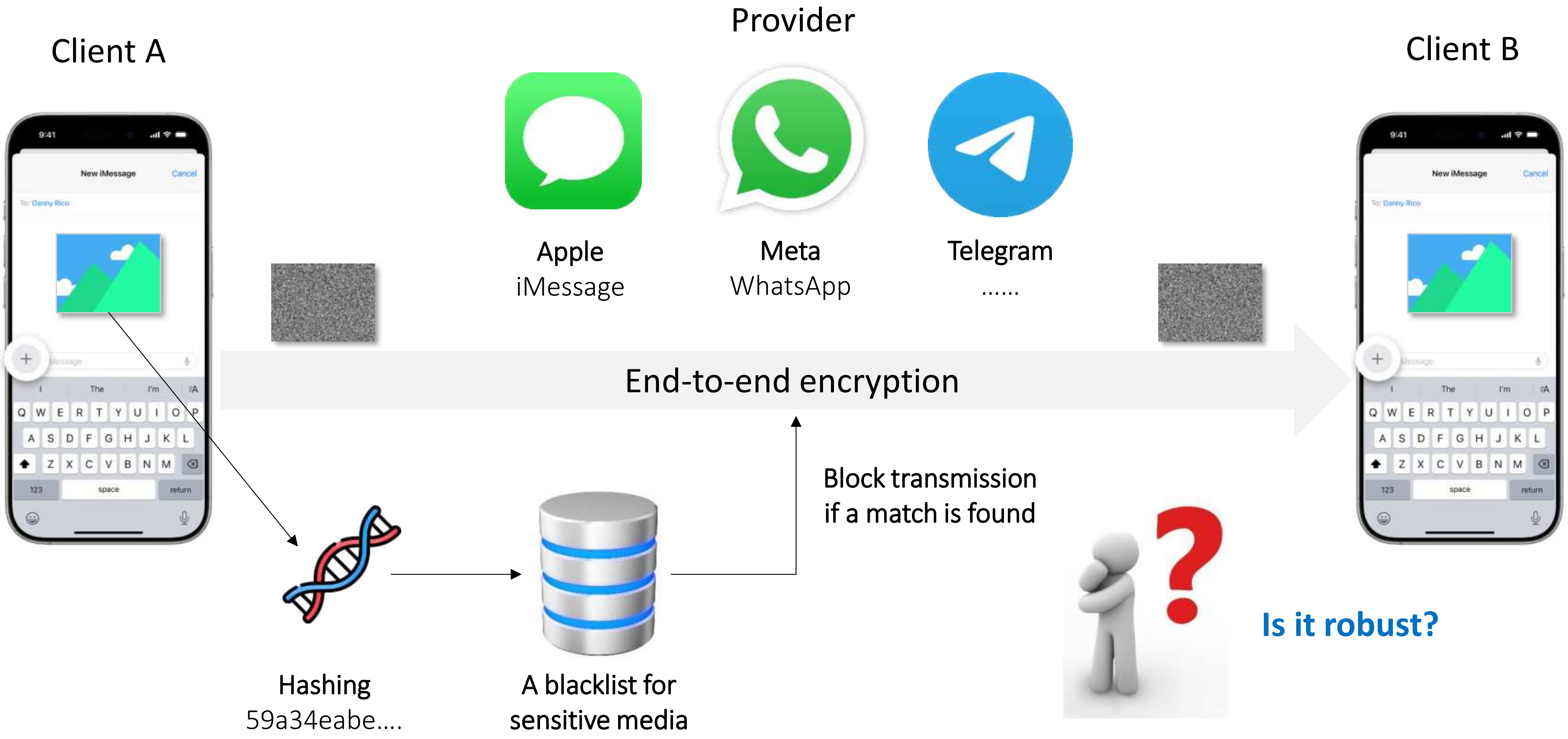
AIGC Watermarking: Robustness and Imperceptibility

	VAE based		DM based	
Method	Bmshj'18	Cheng'20	SDv2.1	**Average**
Pixel-level				*0.165*
DwtDct	0.005	0.002	0.003	0.003
DwtDctSvd	0.103	0.124	0.230	0.152
RivaGAN	0.014	0.017	0.123	0.051
Stable Signature	0.541	0.813	0.003	0.452
Content-level				*0.987*
Tree Ring	0.976	0.993	0.943	0.971
Gaussian Shading	1.000	1.000	1.000	1.000
Ours	0.990	0.983	1.000	0.991

	Metrics		
Method	SSIM↑	LPIPS↓	WO-FID↓
Tree Ring	0.47	0.50	43.81
Gaussian Shading	0.20	0.74	48.32
Ours ($\alpha_2 = 0.02$)	**0.75**	**0.20**	**26.50**
Ours ($\alpha_2 = 0.04$)	0.62*	0.31*	35.02*



AIGC Hashing: Motivations



- Y. Zhang, Y. Sun, S. Qi*, et al. Atkscopes: Multiresolution Adversarial Perturbation as a Unified Attack on Perceptual Hashing and Beyond. *USENIX Security*, 2025.

AIGC Hashing: Ideas

Definition 1. ***(Multiresolution perturbation).*** *The addition of multiresolution perturbation is defined as follows:*

$$X'_{(x,y)\in D_{uvw}} = \mathcal{F}^{-1}(\mathcal{F}(X)+\delta), \quad (3)$$

with notations of

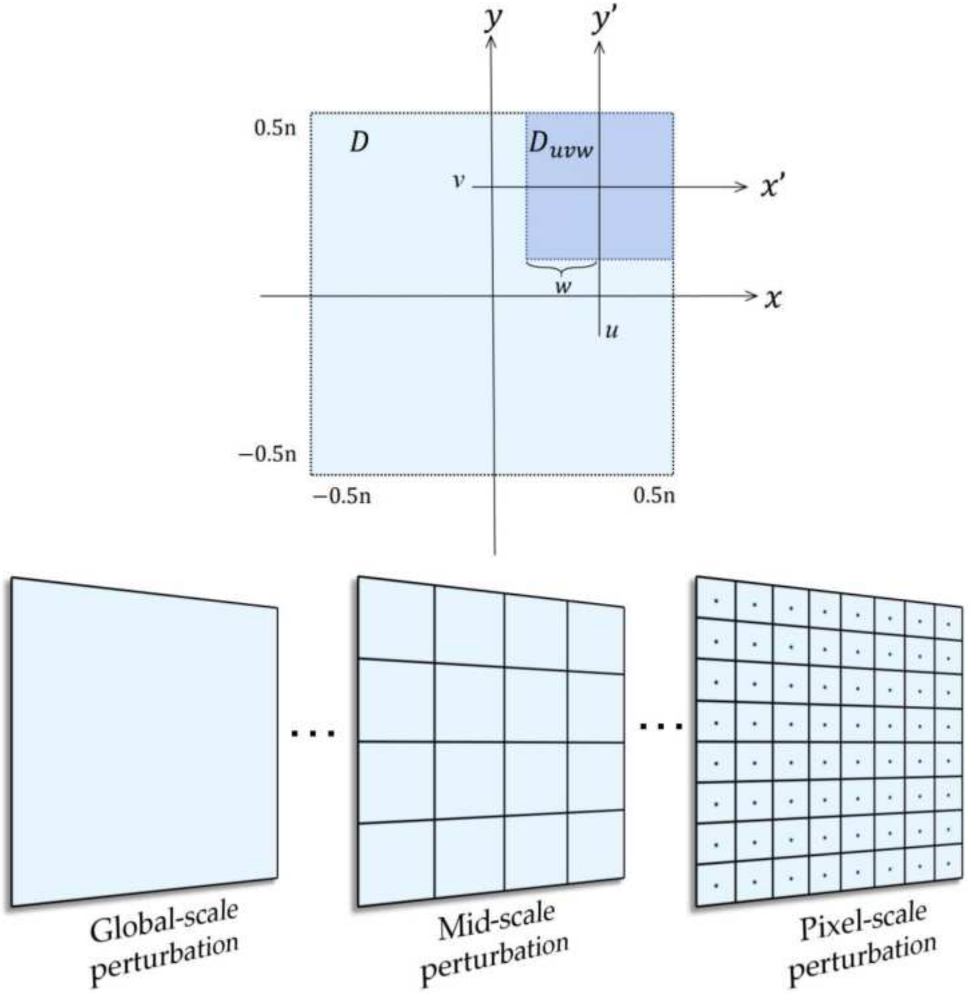
$$\mathcal{F}(X) = \langle X, V_{nm}^{uvw}\rangle = \iint_D (V_{nm}^{uvw}(x,y))^* X(x,y)dxdy, \quad (4)$$

and

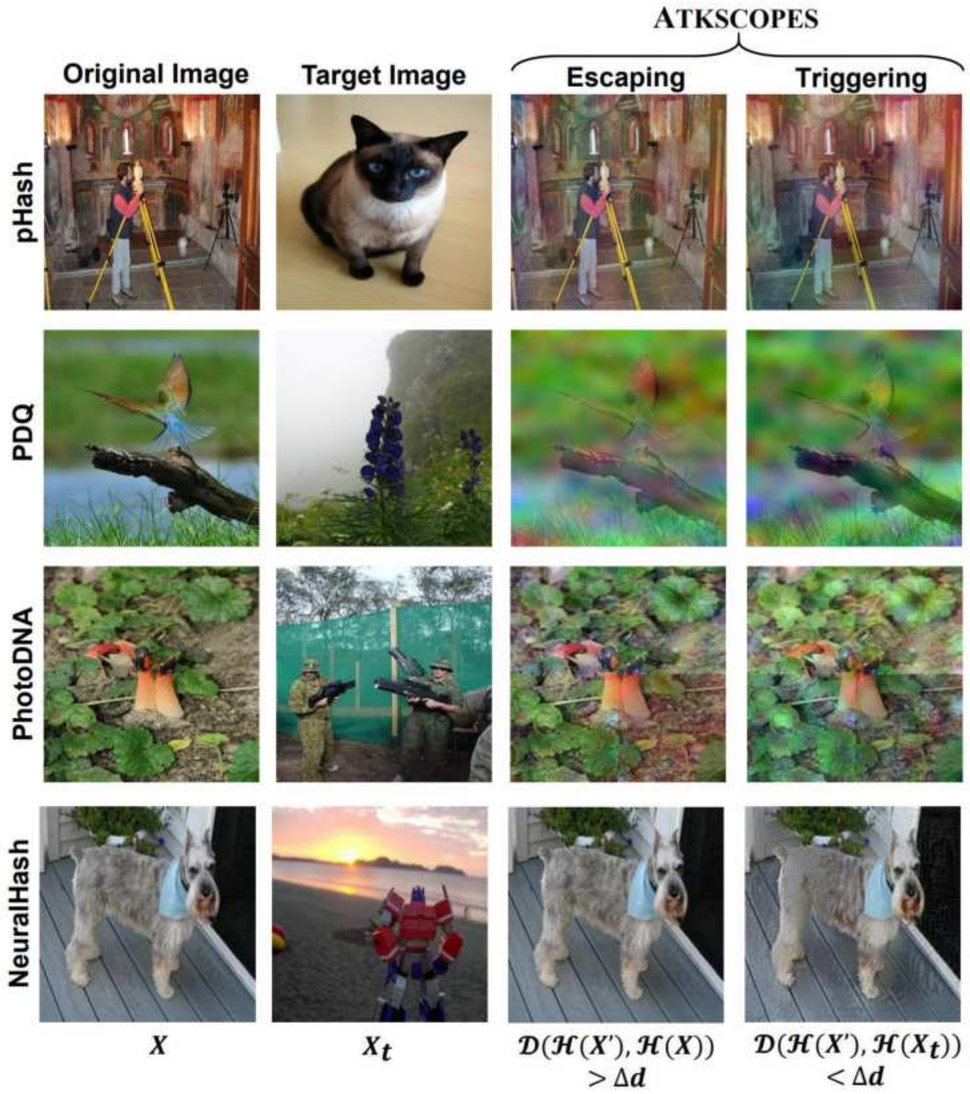
$$\mathcal{F}^{-1}(\mathcal{F}(X)) = \sum_{n,m} V_{nm}^{uvw}(x,y)\mathcal{F}(X), \quad (5)$$

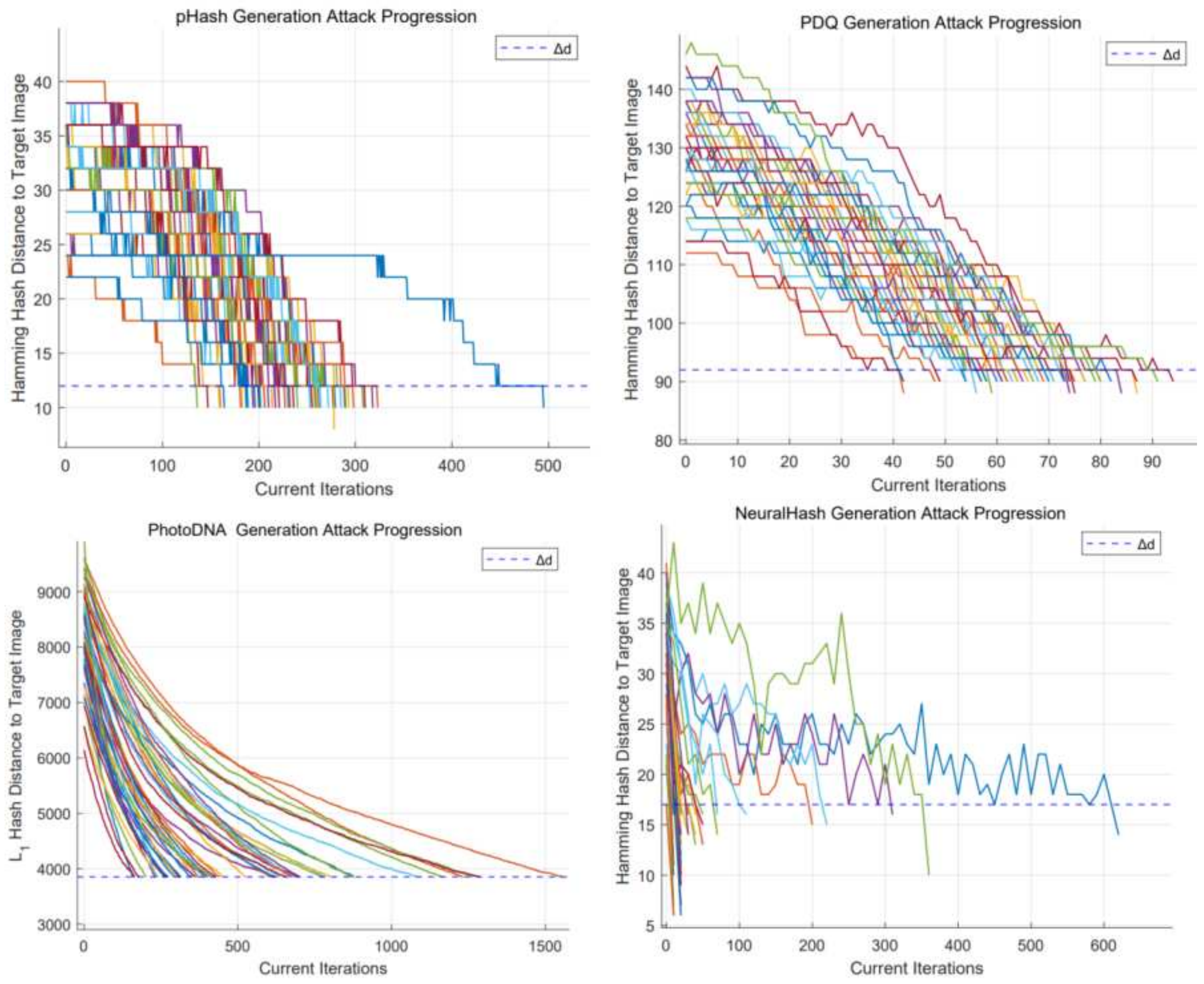
where $\mathcal{F}$ denotes the local orthogonal transformation [39], with image $X(x,y)$ on domain $(x,y)\in D$. The local orthogonal basis function V_{nm}^{uvw} is defined on the domain D_{uvw} with the order parameters $(n,m)\in\mathbb{Z}^2$, converting D to D_{uvw} by the translation offset (u,v) and the scaling factor w, as illustrated in Figure 2. Note that the local orthogonal basis function V_{nm}^{uvw} can be defined from any global orthogonal basis function V_{nm}, e.g., a family of harmonic functions, with following form:

$$V_{nm}^{uvw}(x,y) = V_{nm}(x',y') = V_{nm}\left(\frac{x-u}{w}, \frac{y-v}{w}\right). \quad (6)$$



AIGC Hashing: Uniform, Fast, and Successful Attacks





Conclusion: Our Works for Invariance

Trustworthy AI as background

Symmetry priors in the natural world as principles

Expanding invariant representations at theoretical and practical levels

Invariant Representation

Invariant Theory
CSUR’23

- Moment Invariants → **Refining global invariants** TPAMI’24
- Multiscale and Wavelet → **Designing local invariants** TPAMI’23
- Geometric Deep Learning → **Exploring hierarchical invariants** ICCV’25

Forensic Tasks

AIGC Detection
TIP major

AIGC Watermarking
TIP review

AIGC Hashing
USENIX Security’25

§ Thank you!

by Shuren Qi
shurenqi@cuhk.edu.hk | shurenqi.github.io