High-Performance Self-Supervised Learning by Joint Training of Flow Matching

Kosuke Ukita , Tsuyoshi Okita

Kyushu Institute of Technology

Kyutech
Kyushu Institute of Technology

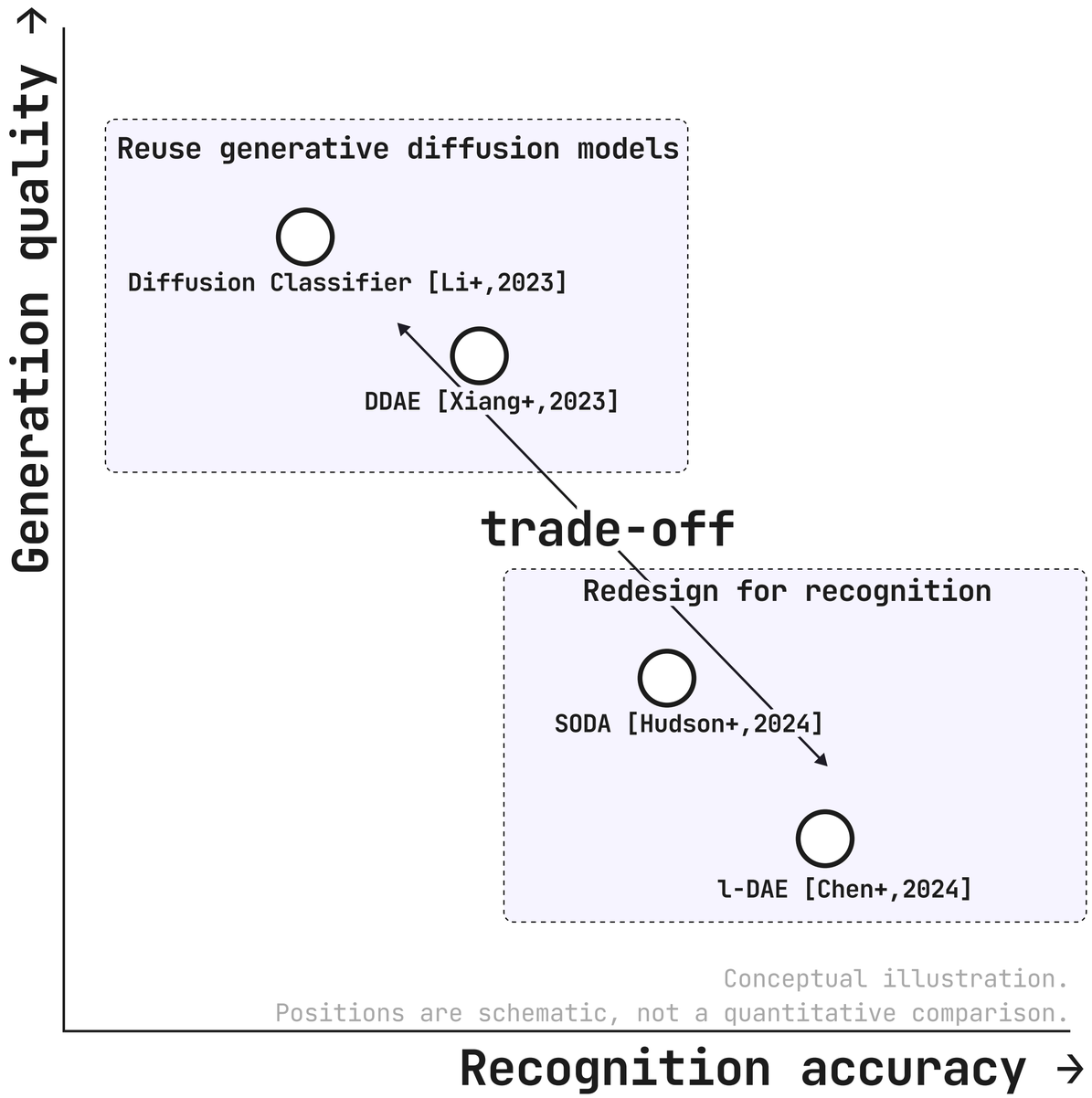
Problem

Great Features.
But a painful trade-off.

Diffusion models learn rich features while generating data - a natural fit for SSL.

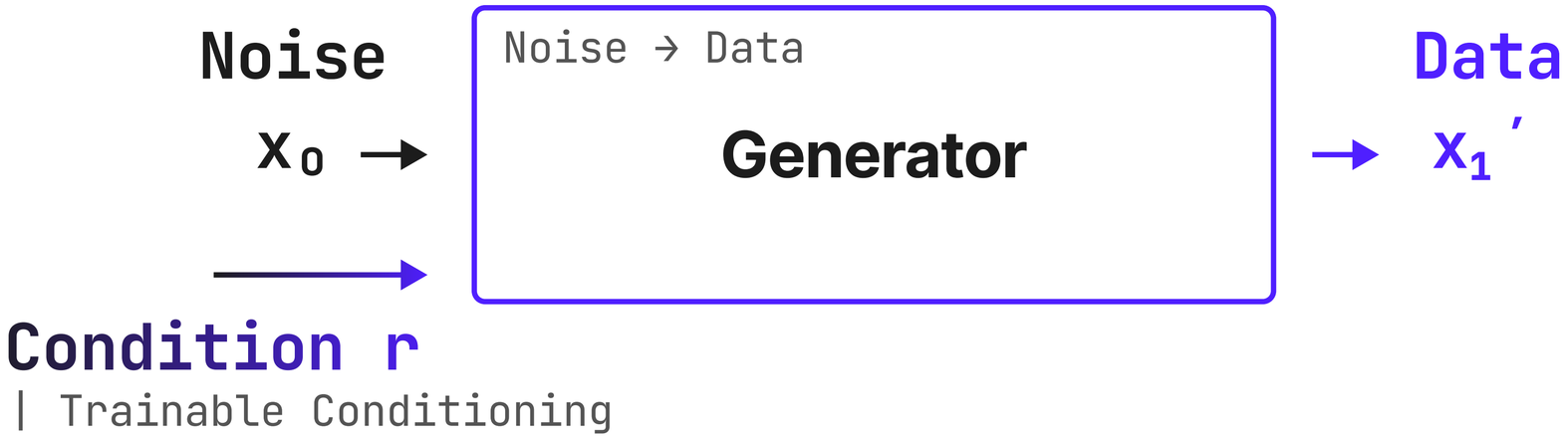
- Keep **generation quality**
  →features are **weak for recognition.**
  [Yang+,2023]

- Tune **for recognition**
  →**generation quality collapses.**
  [Dombrowski+,2024]



Diffusion models are entering SSL

Focus on an overlooked mechanism:
→ The **conditioning** process.

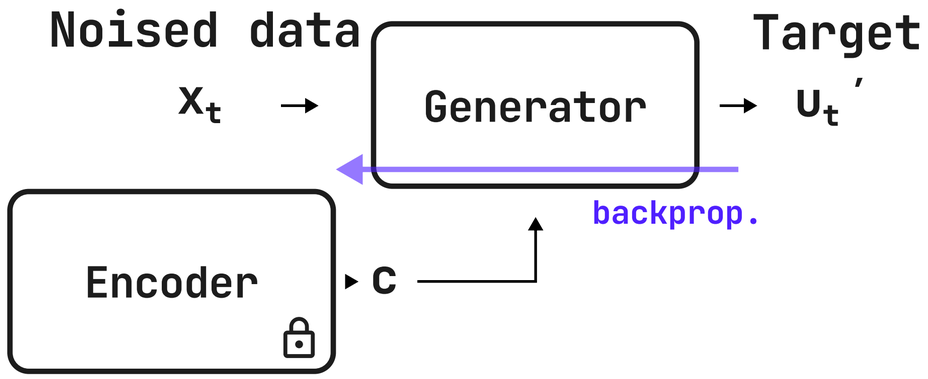


The key idea

Existing

Fixed encoder

Pre-trained, frozen.
Conditions only point the generator toward an intended output.

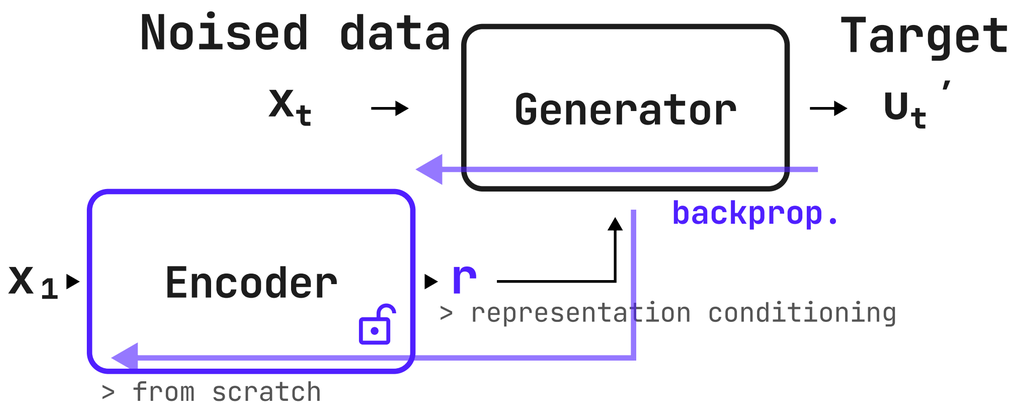


$$\mathcal{L}(\theta) = \mathbb{E}_{t,x_0,x_1}\left[\| v_\theta(x_t, t, c) - (x_1 - x_0) \|^2\right]$$

Ours

Trainable encoder

Trained by the generation loss itself.
From Scratch.

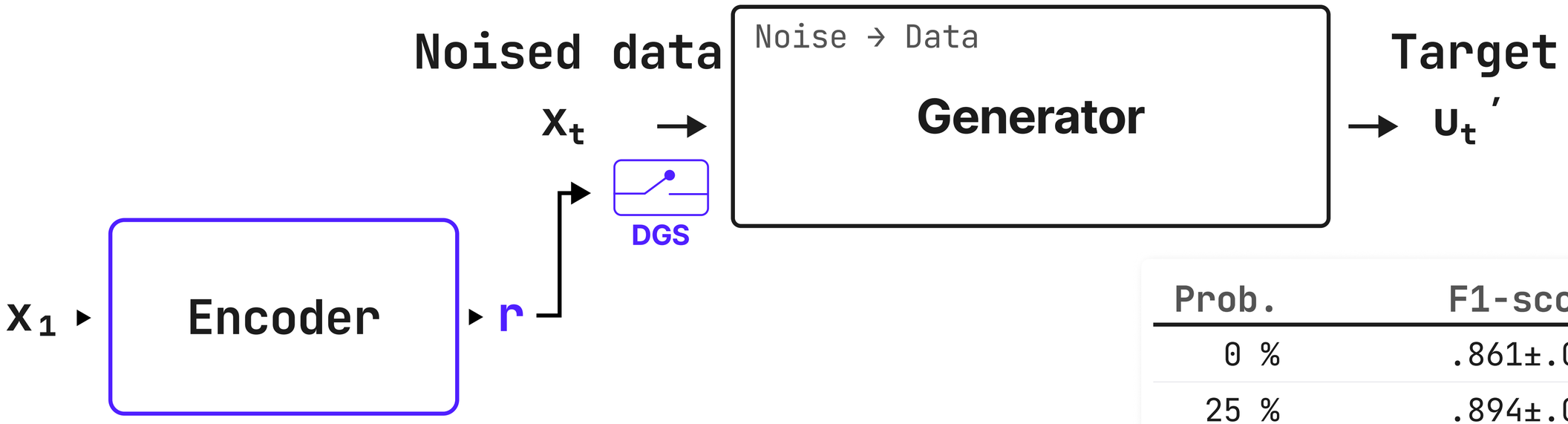


$$\mathcal{L}_{SenFlow}(\theta, \phi) = \mathbb{E}_{t,x_0,x_1}\left[\| v_\theta(x_t, t, f_\phi(x_1)) - (x_1 - x_0) \|^2\right]$$

Dynamic Guidance Switching (DGS)

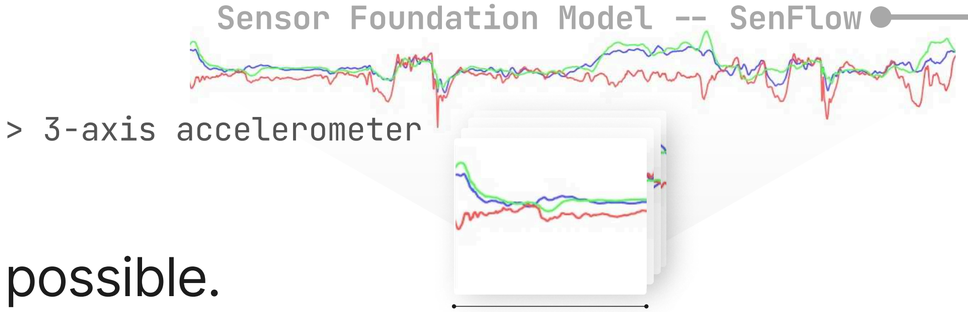
Random masking

50% of the time, the representation is hidden.
Forces the encoder to compress only the essential.



Prob.	F1-score
0 %	.861±.012
25 %	.894±.091
50 %	.924±.073
75 %	.895±.019

> Human Activity Recognition · ADL dataset



Sensor foundation model

Built on wearable sensor data. Many downstream tasks are possible.

Sensor Foundation Model

SenFlow

▼

Downstream Task

- Anomaly detection
- Human Activity Recognition
- Forecasting
- Interpolation
- Signal generation

Downstream Task 1

Human Activity Recognition (HAR)

SOTA on all 5 datasets.

Downstream Task 2

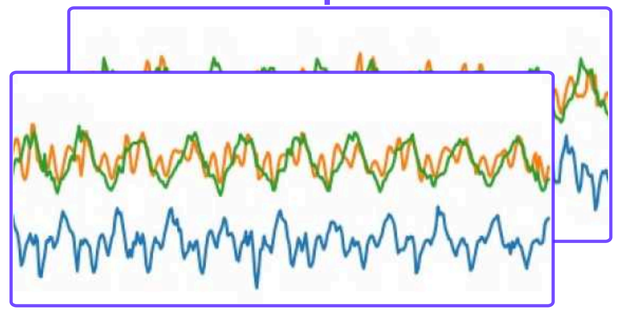
Text-to-Signal generation

Text prompt → realistic sensor signal.
correctly captured.

Prompt

"A young, athletic man, **walking** at a steady and normal pace."

Generated samples



> More at the poster.

Summary

Trainable conditioning

One simple change — powerful representations as a byproduct of generation.

Sensor foundation model

New SOTA on wearable SSL, and generates from text prompts.

Let’s talk at the poster!!

POSTER 195