

# フローマッチングによる異常検知

坂井 俊介<sup>1</sup> and 長谷川 達人<sup>1</sup>

<sup>1</sup> 福井大学大学院

<https://github.com/SkyShunsuke/FlowAD>

## 1 はじめに

教師なし異常検知 (Unsupervised Anomaly Detection; UAD) は、異常画像を用いずに、正常画像のみから欠陥や未知の特徴を含む異常を検出する問題設定である。昨今の深層学習を用いた UAD では、(i) 事前訓練した特徴抽出器の活用と (ii) 特徴空間上での正常特徴の分布モデル化によって著しい発展を遂げてきた [5]。なかでも、深層生成モデルである拡散モデルを用いて、(ii) を実現する手法が高い検出精度を達成できることから注目を集めている [5]。

拡散モデルを用いた UAD のほとんどの手法は、ノイズ除去による再構成誤差によって異常を判別する。これらの手法は高い検出性能を示す一方で、どの程度強い強度のノイズを加えるかは人手で決める必要がある。また、拡散モデルは深層ニューラルネットワークによる反復的なノイズ除去を必要とするため、推論が遅い [5]。こうした問題は、拡散モデルを用いた UAD の実問題への活用を妨げる要因となっている。

これらの問題を解消するため、本研究ではフローマッチング (Flow Matching; FM) を用いた UAD の枠組みを新たに確立する。FM [3] は、生成モデリングの一種であり、ノイズ分布とデータ分布を補間する経路を学習する。

拡散モデルと比較すると、FM のサンプリング軌道は直線的であり、高速にサンプリングできる。加えて、ノイズからデータへの変換は決定論的な過程で表されるため、学習した生成モデル上の尤度を推定できる。以上より、FM は拡散モデルよりも速い推論でより多様な異常検知の手段を提供すると考えられる。しかしながら、FM を用いた UAD においては、以下のような点が未解明である。

- Q1. FM より、どのように異常を検出できるか？
- Q2. 拡散モデルと比較して、速度・精度は優れているか？

## 2 フローマッチングによる異常検知

図 1 に本研究で提案する FM による異常検知の概要を示す。FM は、生成モデリングの一種であり、 $X_1 \sim q$  と正規分布からのノイズ  $X_0 \sim \mathcal{N}(0, I)$  を補間する速度場を深層ニューラルネットワーク  $u_t^\theta \approx u_t$  により学習する (図 1 左)。本研究では、先行研究 [5, 1] と同様に事前訓練した特徴抽出器 (Enc) の特徴空間上で FM を学習する。

FM を用いる UAD においては、正常分布からの乖離度は、(a) 密度推定、(b) 逆推定、(c) 再構成のいずれかによって評価できる (図 1 右)。密度推定は、フローマッチングを用いた最

Flow Matching-based Anomaly Detection, <sup>1</sup> Shunsuke Sakai, Tatsuhito Hasegawa (Graduate School, University of Fukui).

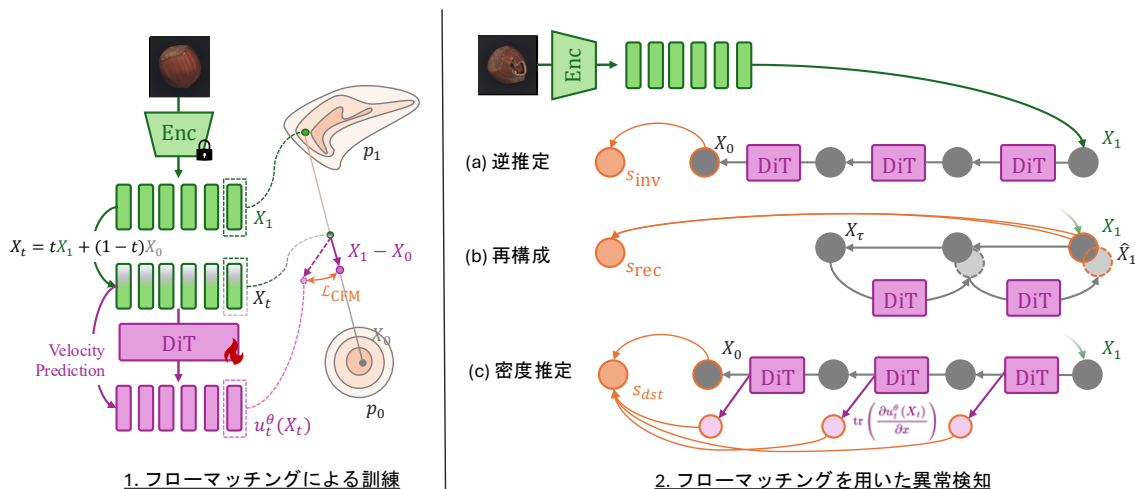


図 1 フローマッチングによる異常検知の概要。FM では、図左のようにデータ・ノイズペアをサンプルし、線形補間した経路上の点から速度を予測するようにニューラルネットワーク (DiT) を訓練する。図右に、本研究で提案する FM で学習した DiT を用いた 3 種類の検出手法を示す。

表1 検出性能および推論効率の定量評価. 検出性能は MVTecAD における各指標の値の全クラス平均を取って評価する.

Method	Image-level			Pixel-level				Efficiency	
	AUROC	AP	F1max	AUROC	AP	F1max	AUPRO	FPS	Memory [GB]
<b>Non-Diffusion UAD</b>									
HVQ-Trans [4]	98.0	99.5	97.5	97.3	48.2	53.3	91.4	5.6	N/A
CFlowAD [2]	92.7	97.2	94.0	95.8	46.8	49.6	89.0	31	<b>2</b>
<b>Diffusion Model UAD</b>									
InvAD (逆推定) [5]	<b>99.0</b>	99.6	<b>98.5</b>	97.5	46.5	52.3	92.7	<b>41</b>	10
OmiAD (再構成) [1]	98.8	<b>99.7</b>	<b>98.5</b>	<b>97.7</b>	<b>52.6</b>	<b>56.7</b>	93.2	17	N/A
<b>Flow Matching (Ours)</b>									
(a) 密度推定	98.4	99.3	97.9	97.4	45.9	51.6	92.5	5	53
(b) 逆推定	98.4	99.3	97.9	97.4	46.8	52.2	92.8	<b>41</b>	10
(c) 再構成	98.7	99.5	98.3	97.4	50.6	54.4	<b>93.5</b>	<b>41</b>	10

も単純な異常検知のアプローチであり, 以下のように行える:

$$s_{\text{dst}}(X_1) := \log p_0(X_0) - \int_0^1 \text{tr} \left( \frac{\partial u_t^\theta(X_t)}{\partial x} \right) dt, \quad (1)$$

低い密度を持つ点を異常として定義しているため, 異常度は負の対数尤度として定義される. 逆推定は  $u_t^\theta$  を  $t: 1 \rightarrow 0$  に積分し, 与えられた正常画像  $X_1$  を生成するようなノイズ点  $X_0$  を推定する. 訓練時に遭遇していない異常箇所については高ノルムなノイズ点  $X_0$  に写像されるため, 推定後のノルムを評価することで異常を検出できる:

$$s_{\text{inv}}(X_1) := \|X_1 - \int_0^1 u_t^\theta(X_t) dt\|_2^2. \quad (2)$$

逆推定は, 式1の右辺第一項のみを考慮したものであり, 密度推定の大幅な近似である. 再構成は, 入力画像  $X_1$  とノイズ  $\epsilon$  の線形経路上の点  $X_\tau$  から元の点  $X_1$  を復元した際の再構成誤差によって異常を検出する:

$$s_{\text{rec}}(X_1) := \|X_1 - (X_\tau + \int_\tau^1 u_t^\theta(X_t) dt)\|_2^2, \quad (3)$$

ここで,  $X_\tau = tX_1 + (1-t)\epsilon$  である. 再構成は拡散モデルと同様に速度場  $u_t^\theta$  によるノイズ除去を介した異常検知である.

### 3 実験

本研究では, FM を用いた UAD において, これらの3種類の異常度推定手法を確立し, その性能を検証することで Q1/2 を検証する. 全ての実験は15種の産業画像異常検知タスクを含む MVTecAD において実施し, 全クラスの平均性能を報告する. 検出性能は画像・画素単位の指標を用いてそれぞれ評価し, 推論効率は1秒あたりの処理画像枚数 (FPS) および推論時の VRAM 使用量 (Memory) を評価する<sup>\*1</sup>. 図1の特徴抽出器としては EfficientNetB4 を, 速度場  $u_t^\theta$  として Diffusion Transformer (DiT) を用い, MVTecAD の全クラスに対して単一の DiT を学習する. 速度場の数値積分を解く際にはオイラー法を用いる.

表1に異なる UAD の枠組みにおける検出性能と推論効率の定量評価結果を示す. まず, FM の異なる異常検知のアプローチ (a-c) を比較した場合, 検出性能・推論効率ともに再構

成が最も優れていることが分かった. 拡散モデルを用いる場合は再構成ステップ数を大きく取る必要があったが, FM ではサンプリング軌道が直線的であるため, 少ない再構成ステップでも高い検出性能を達成できる. 逆推定は検出性能では再構成に劣るが, どこまで摂動させるかのハイパーパラメータである  $\tau$  の設計が不要である. また, 密度推定は推論効率がその他の手法と比較して顕著に悪く, 性能も劣る. 以上より, Q1 については, “FM を用いた UAD においては, 逆推定もしくは再構成が検出性能・推論効率の観点から優れている” と言える. 一方で, 拡散モデルによる UAD [1, 5] と比較すると, 提案手法のいずれもその性能を上回ることができていない. 以上より, Q2 については, “匹敵する検出性能・推論効率を示すが, 拡散モデルを用いた UAD を上回ることには現時点ではできていない” と結論づけられる.

### 4 おわりに

本研究では, FM を用いた新たな UAD の枠組みを確立し, 異なる検出手法の有効性を検出性能・推論効率の観点から検証した. また, 拡散モデルと FM を用いた UAD と比較した場合, 依然として拡散モデルが優れており, FM を用いた UAD には未だ改善の余地が残る. 拡散モデルと比較すると, FM は補間経路や事前分布の設計の自由度が高い. 今後は UAD における補間経路や事前分布の設計により, 検出性能や推論効率の改善を図る.

### 謝辞

本研究の一部は, JSPS 科学研究費助成事業基盤研究 A(25H01110) 及び基盤研究 C(23K11164) の助成によるものである. ここに謝意を表す.

### 参考文献

- [1] Y. Feng et al. “OmiAD: One-Step Adaptive Masked Diffusion Model for Multi-class Anomaly Detection via Adversarial Distillation”. *ICML*. 2025.
- [2] D. Gudovskiy et al. “CFlow-AD: Real-Time Unsupervised Anomaly Detection With Localization via Conditional Normalizing Flows”. *WACV*. 2022.
- [3] Y. Lipman et al. “Flow Matching for Generative Modeling”. *ICLR*. 2023.
- [4] R. Lu et al. “Hierarchical Vector Quantized Transformer for Multi-class Unsupervised Anomaly Detection”. *NeurIPS*. 2023.
- [5] S. Sakai et al. “InvAD: Inversion-based Reconstruction-Free Anomaly Detection with Diffusion Models”. 2025.

<sup>\*1</sup> 全ての実験は NVIDIA H100 1機を用いて実施した.