

速度場を通じた物理的整合性の取れた群衆流動予測

川端 智紀[†] 落合 桂一[†] 戸田 浩之[†]

[†] 横浜市立大学大学院データサイエンス研究科 〒236-0027 神奈川県横浜市金沢区瀬戸 22-2

E-mail: †{y245604c,ochiai.kei.dk,toda.hir.xg}@yokohama-cu.ac.jp

あらまし 都市において、都市内の各エリアに関する未来の群衆流動（例：群衆量，流入，流出量）を正確に予測することは大きな利益をもたらす。そのため、群衆流動を予測するためのモデル，特に深層学習ベースのモデルが多く提案されている。しかし，深層学習ベースを含む既存モデルは，群衆流動に内在する本質的な物理法則を明示的に考慮していない。具体的には，あるエリアの群衆量の増減は流入量と流出量の差分に一致するという関係性を明示的に考慮していない。そのため，既存モデルは，信頼性の低下につながる物理的に不整合な予測を行う可能性がある。そこで，本研究では，物理的に整合性の取れた群衆流動予測を行うために，推論プロセス自体に物理法則を組み込む予測フレームワークを提案する。提案するフレームワークでは，群衆を移動させるダイナミクスを表す速度場を予測した後，予測した速度場を用いて未来の群衆流動の予測を行う。実世界群衆流動データを用いた実験の結果，物理的整合性の点で，提案フレームワークは既存モデルより優れた予測が可能であることが示された。

キーワード 移動・行動分析，モビリティ，時空間予測，PIML，スマートシティ

1 はじめに

群衆流動予測とは，都市広域を複数のグリッドに分割し，データに基づいて未来の各グリッドの群衆集中度や流入流出量を予測するタスクである。これらの値を正確に予測することは，都市のステークホルダーに対して大きな利益をもたらす。例えば，都市管理者の観点では，事前に群衆事故のリスクを評価することが可能になるため，適切な警備計画の策定を行うことができるようになる。また，小売業者の観点では，特定のエリア・時間帯における潜在的な顧客数を把握することが可能になるため，より正確な需要予測を行うことができるようになる。

このような社会的背景から，群衆流動予測をグリッドベースの時空間予測問題として扱い，取り組む研究が多く存在する [1]。特に，群衆流動に内在する複雑な時空間パターンを捉えるために，深層学習ベースの予測モデルが多く提案されている。これらの既存モデルは，グリッド単位の予測値と真値のずれを平均二乗誤差（MSE）や絶対平均誤差（MAE）などの評価指標で測ると，統計ベースのモデル（例：ARIMA）を上回る予測精度を示している。

一方で，これらの既存モデルの多くは，群衆流動データを観測する現実世界の物理法則を十分に考慮していない。さらに，図1上部のように，群衆量データと流入量・流出量データの2種類の群衆流動データが存在する場合も，個別にモデルを学習し予測を行っている。具体的には，一種の質量保存則とみなすことができる，あるエリアの群衆量の増減は流入量と流出量の差分に一致するという関係性を無視した予測を行う可能性がある。すなわち，流入量・流出量用の予測モデルが導いた未来の流入量・流出量に従い，過去の群衆量を変化させた未来の群衆量の予測結果は，群衆量用の予測モデルが導いた予測結果と常に一致するとは限らない。この物理的に不整合な予測は，予測

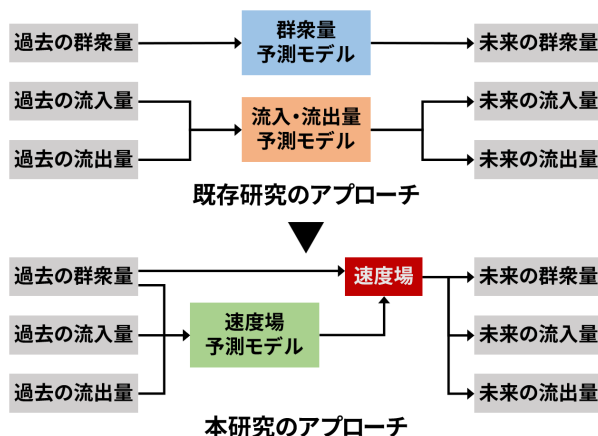


図1 本研究でとるアプローチの概要。

結果を解釈・利用する者の意思決定に混乱を招き，信頼性を著しく低下させる恐れがある。例えば，群衆量予測モデルは混雑が緩和することを予測しているのに対して，流入量・流出量予測モデルはさらに流入することを予測している場合，都市管理者は警備員を減員すべきか増員すべきか判断することが難しくなる。

また，この物理的に不整合な予測を行う問題は，既存モデルを群衆量・流入量・流出量の同時予測を行うマルチタスク学習へ拡張し，質量保存則に関する損失項を追加することでは解決できない。なぜならば，前述の質量保存則は，あるグリッドでは群衆を消失させ，あるグリッドで群衆を湧出させるという形でも満たすことができるためである。すなわち，物理的に無意味な形で質量保存則を満たすように学習する可能性があるためである。また，損失を考慮することができるのは，あくまで学習時のみであり，推論時に推論時に質量保存則を満たすように誘導する効果はないためである。

そこで、本研究では、物理的に整合性の取れた群衆流動予測を行うために、推論プロセス自体に物理法則を組み込む予測フレームワークを提案する。提案するフレームワークでは、未来のタイムステップにおける群衆量・流入量・流出量を直接予測するのではなく、図1下部のように物理法則に従って群衆を移動させるダイナミクスを予測する。具体的には、ある時刻に、ある空間上の位置にいた量が、どの方向にどれだけの速さで動くかを表す、速度場を予測する。そして、予測した速度場に従い過去の群衆量を変化させ、未来の群衆量・流入量・流出量の予測を行う。速度場を通じた予測を行うことで、物理的に意味のある形でのみ学習し、推論時も物理法則を考慮した予測を行うことができるようになる。

2 関連研究

2.1 群衆流動予測に関する研究

群衆流動予測を含む時空間予測において、時空間依存関係を捉えることが高精度な予測のために必須である。そのため、この分野では、時間的依存関係を捉えることができるモジュールと空間的依存関係を捉えることができるモジュールを融合した、深層学習ベースの時空間予測モデルが多く提案されている。時空間予測モデルに関する代表的な手法として、Convolutional LSTM (ConvLSTM) [2] が存在する。降水予測のために提案されたモデルであるが、交通予測において統計ベースのARIMAモデルやシンプルなLSTMを上回る予測精度を示すことが実証されている [3]。そして、ConvLSTM以降、LSTMやCNNよりさらに高度な手法を融合した予測モデルが多く提案されている。例えば、近年では、都市広域にまたがる時空間的な依存関係を捉えるために、ConvLSTMをピラミッド型に積層するモデル [4] や、マルチスケール時空間依存関係を表現可能なメモリユニットを導入したモデル [5] が存在する。一方で、これらのモデルは優れた予測精度を実現しているが、現実世界の物理法則を十分に考慮していない。特に、あるエリアの群衆量の増減は流入量と流出量の差分に一致するという関係性は、群衆流動に内在する本質的な法則であるにもかかわらず、この関係性を十分に考慮した研究は私の知る限り存在しない。

2.2 物理法則を組み込んだ機械学習手法に関する研究

従来のデータドリブンな機械学習モデルに、物理法則や事前知識を融合する分野として Physics-Informed Machine Learning (PIML) が存在する [6] [7]。物理法則を融合することで、機械学習モデルを物理的に尤もらしい解へと導くことが可能となる。

この分野における代表的なアプローチが、Physics-Informed Neural Network (PINN) である。標準的なPINNは、主に偏微分方程式 (PDE) や常微分方程式 (ODE) に関する問題を解くための枠組みであり、ニューラルネットワークの損失関数に微分方程式とその条件を組み込むことで、微分方程式の解を近似する [8]。すなわち、PINNは、微分方程式で記述される物理法則をニューラルネットワークの学習プロセスに統合し、物理法則を満たす解を探索することができる。また、微分方程式

で記述される法則であれば、その法則を損失関数に利用できるというPINNの特徴に着目し、微分方程式が支配方程式として機能している分野でPINNの活用が進んでいる。例えば、疫学分野では、感染症のダイナミクスを記述するSIRモデルを組み込んだPINNが、ウイルスの拡散傾向を正確に特定・予測できることが示されている [9]。

また、観測ノイズなどにより物理法則を満たさないデータを用いて、物理法則を組み込んだモデルを学習しようとする、学習の停滞や不安定な学習を引き起こす可能性がある。これは、モデルは物理法則を満たす解空間しか探索できないのに対し、ノイズを含むデータとの誤差を最小とする解はこの解空間の外に存在する。この問題を解決するために、モデルを学習する前に、観測データを物理法則を満たすように前処理を行うことがある。前処理の手法として、最適化ベースの射影がある [10] [11]。具体的には、満たすべき物理法則を制約条件とし、制約条件を満たし観測値に最も近いデータポイントを探索する最適化問題を解くことで前処理を行う。この最適化問題を解くことは、ノイズにより物理法則を満たしていないデータを、物理法則を満たす多様体上の最も近い点に射影することを意味する。

3 準備

3.1 問題設定

既存研究 [4] [5] では、直近 (Closeness) の観測値、1日前 (Period) の観測値、1週間前 (Trend) の観測値およびメタ情報を用いて、次の時刻の観測値を予測する。直近の観測値、1日前の観測値、1週間前の観測値として、それぞれ L タイムステップ分の情報を用いる場合、これらは以下のように表記できる：

$$\begin{aligned} \mathbf{X}^{\text{Closeness}} &= [\mathbf{X}_{t-L+1}, \mathbf{X}_{t-L+2}, \dots, \mathbf{X}_t] \\ \mathbf{X}^{\text{Period}} &= [\mathbf{X}_{t-L+2-l_P}, \mathbf{X}_{t-L+3-l_P}, \dots, \mathbf{X}_{t+1-l_P}] \\ \mathbf{X}^{\text{Trend}} &= [\mathbf{X}_{t-L+2-l_T}, \mathbf{X}_{t-L+3-l_T}, \dots, \mathbf{X}_{t+1-l_T}]. \end{aligned} \quad (1)$$

ここで、 l_P, l_T は、それぞれ1日間の観測回数と1週間の観測回数である。また、メタ情報として、観測時刻に関する時刻情報、曜日情報、休日・祝日情報 \mathbf{v}^{meta} を活用する。これらの過去の観測値 $\mathbf{X}^{\text{Closeness}}, \mathbf{X}^{\text{Period}}, \mathbf{X}^{\text{Trend}}$ と対応するメタ情報 \mathbf{v}^{meta} を用いて、次の時刻の観測値 \mathbf{X}_{t+1} を予測することは、数学的には以下の式で定式化される：

$$[\mathbf{X}^{\text{Closeness}}, \mathbf{X}^{\text{Period}}, \mathbf{X}^{\text{Trend}}, \mathbf{v}^{\text{meta}}] \xrightarrow{f} \hat{\mathbf{X}}_{t+1}. \quad (2)$$

ここで、 f は群衆流動予測を行う関数であり、学習される予測モデルである。この問題定義の特徴は、直近の観測値だけでなく1日前・1週間前の観測値を、予測に活用することである。そのため、1週間分の長期の観測値のすべてを用いることなく、群衆流動に内在する周期性を明示的に活用した効果的な予測が可能となる。

本研究でも、上記の問題設定は踏襲する。ただし、物理的に

不整合な予測を行う問題を解決するために、群衆量と流入量・流出量の間には存在する関係を明示的に考慮する。具体的には、以下の式 (3) で表される、群衆量 **Density** と流入量 **Inflow**・流出量 **Outflow** の間に存在する関係を制約として用いることで明示的に考慮する：

$$\text{Density}_{t+1} - \text{Density}_t = \text{Inflow}_{t+1} - \text{Outflow}_{t+1}. \quad (3)$$

この式 (3) は、観測ノイズが存在しないとき、群衆量の増減は流入量と流出量の差分に一致するという関係が各グリッドで成立することを意味する。

3.2 速度場の基本的知識

速度場および関連する密度・流束という概念について概説する。これらは、流体や確率密度において重要な概念であり、ある空間上の位置 x と時間 t を用いて定義される。

- **速度場 $\mathbf{v}(x, t)$** ：時刻 t において、位置 x にいた量（例：質量、電荷、確率）の速度を表すベクトル値関数である。すなわち、時刻 t に位置 x にいた量が、どの方向にどれだけの速さで動くかを表す。
- **密度 $\rho(x, t)$** ：時刻 t において、位置 x 周辺の単位体積に含まれる量を表すスカラー量である。
- **流束 $\mathbf{j}(x, t)$** ：時刻 t において、位置 x 周辺の単位時間あたり単位面積を通過する量を表すベクトルである。これは、以下の式 (4) のように密度と速度場の積で求めることができる

$$\mathbf{j}(x, t) = \rho(x, t)\mathbf{v}(x, t). \quad (4)$$

また、これらは質量保存則を示す連続の式

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \quad (5)$$

で関連付けられる。ここで、 ∇ は空間変数に対する微分作用素である。連続の式 (5) とは、この流束の発散と密度の時間変化の関係を表した式である。すなわち、密度の変化は、流出量と流入量のみによって決定されるという関係を表した式である。

これらの概念は、群衆流動に対しても適用可能である。単位面積をグリッドとした場合、グリッド内の群衆量は密度に、流入量・流出量は流束によりグリッド境界を横断した群衆量に一致する。そして、速度場は群衆のダイナミクスを表す関数となる。そのため、現在の群衆量の分布を速度場により移動させた結果として、未来の群衆量の予測結果が得られる。この際、各グリッドに関して、グリッド境界を横断する群衆の量を集計することで、未来の流入量・流出量の予測結果が得られる。また、連続の式 (5) は、群衆が忽然と現れたり消えたりしない限り、群衆量の変化は流入量・流出量のみによって決定されるという関係を表す式となる。

4 提案フレームワーク

本節では、本研究で提案するフレームワーク、すなわち物理的整合性の取れた群衆量・流入量・流出量予測を行うための群

衆流動予測フレームワークについて詳述する。まず、提案フレームワークの概要について説明した後、アーキテクチャについてモジュールごとに説明する。その後、学習戦略について説明する。

4.1 提案モデルの概要

前節で述べたように、速度場を現在の群衆量の分布に対して適用することで、物理的整合性の取れた未来の群衆量・流入量・流出量の予測が可能となる。そのため、提案モデルは、未来の群衆量・流入量・流出量を直接予測するのではなく、速度場を予測する。

一方で、1つの速度場が記述できるのは、ある地点における単一の平均的なダイナミクスのみである。しかし、ある地点の群衆は、さまざまな属性や目的を持つ集団が混ざり合い出来たものである。そのため、このダイナミクスを単一の速度場のみで記述しようとすると問題が生じる。例えば、ある駅から東に行く群衆と西に行く群衆が同程度存在する状況を考える。このとき、単一の速度場で記述できるのは単一の平均的なダイナミクスのみであるため、東西への移動は存在しないと学習してしまう。よって、群衆のダイナミクスを速度場で学習するためには、単一の速度場のみでは不十分であり、複数の速度場が必要となる。

以上のことから、提案モデルは、現在の群衆量の分布を移動させるための、複数の速度場の予測を行う。具体的には、U-Netベースのアーキテクチャを用いて、過去の群衆量・流入量・流出量の観測値とメタ情報から、複数の速度場と各グリッドにおける速度場に対する重みを予測する。各グリッドにおいて速度場に対する重みも予測することで、グリッドごとに複数の速度場を動的に混合することが可能となる。すなわち、群衆の複雑なダイナミクスを、単純な動きに動的に分離して予測を行うことが可能となる。

学習においても、複数の速度場で群衆の複雑なダイナミクスを学習できるように誘導する。具体的には、速度場ごとに異なるダイナミクスを学習するように、多様性損失やサイクル一貫性損失 [12]・直交性制約を学習時に考慮する。また、提案モデルは、単一モデルで群衆量予測・流入量予測・流出量予測という複数のタスクを解くように学習する。そのため、通常の学習戦略では、タスク間で競合が発生し、いずれかのタスクの予測精度が犠牲になる可能性がある。そこで、全タスクのバランスをとるパレート改善を目指す学習戦略を導入する。

4.2 アーキテクチャ

本節では、図 2 に示す提案するフレームワークのアーキテクチャについて詳述する。

まず、提案フレームワークを構成する各モジュールについて説明する前に、モデルの全体像を説明する。提案フレームワークは、主に 3つの段階に分かれている。

1. **準備段階**：この段階では、過去の群衆流動データおよび対応するメタ情報から、速度場予測を強化するための条件を生成する。具体的には、軽量の**マクロ群衆量予測器**を用い

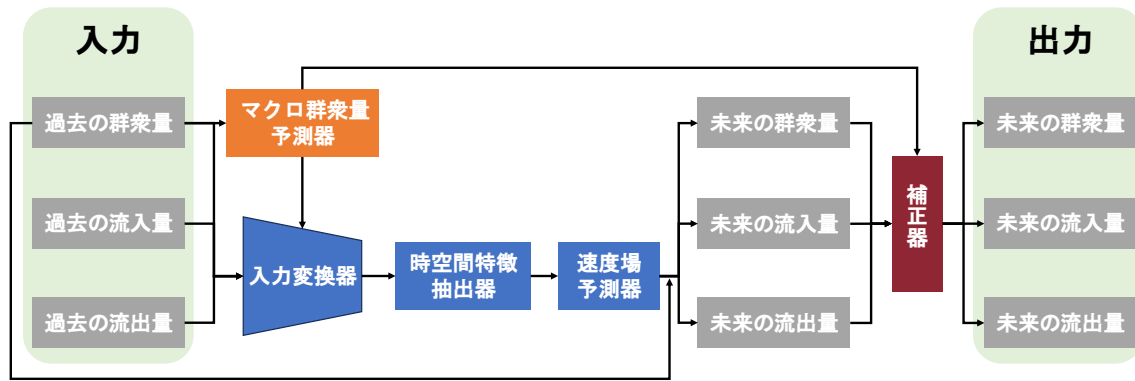


図2 提案するフレームワークの全体像。

て、いくつかのグリッドがまとまったマクロな領域の未来の群衆量を予測する。この未来のマクロな群衆量は、後続の速度場予測を行うモジュールにおいて、強力な条件として用いることができる。具体的には、速度場予測モジュールは、未来にどこに人がいるべきかを情報として活用することができるため、ゼロから未来を予測する必要がなくなり、学習の難易度を大きく下げることができる。

2. **予測段階**：この段階では、過去の群衆流動データおよび対応するメタ情報、準備段階で予測したマクロ群衆量を用いて、現在の群衆量分布をどのように移動させるべきかという速度場を予測する。具体的には、まず**入力変換器**を用いて、過去の群衆流動データ・メタ情報・マクロ群衆量を統合し、後続のモジュールで活用する豊かな初期表現を形成する。その後、U-Net ベースの**時空間特徴抽出器**を用いて、得られた初期表現から時空間的な特徴を抽出し、速度場予測に適した高次元空間へ射影する。最後に、**速度場予測器**を用いて、高次元空間の特徴をもとに、具体的な複数の速度場と各速度場に対する重みを予測する。そして、予測した速度場および重みに従い、未来の群衆量・流入量・流出量の予測値が得られる。
3. **補正段階**：この段階では、群衆量・流入量・流出量の予測値を補正することである。具体的には、**補正器**を用いて、マクロ群衆量予測器により予測された未来のマクロ群衆量と整合性が取れるように、群衆量の予測値を補正する。さらに、群衆量のみ補正した場合、物理的な整合性が取れなくなるため、流入量・流出量に対しても補正を行う。ただし、この段階は推論時のみ適用する。

以降の小節では、提案フレームワーク内の各モジュールについて説明する。

4.2.1 マクロ群衆量予測器

このモジュールの目的は、速度場を予測するモジュールを強化するための条件を生成することである。具体的には、未来の各グリッドのミクロな群衆量を予測するのではなく、複数グリッドの群衆量を合計した未来のマクロな群衆量を予測する。未来のマクロ群衆量を入力および損失計算の対象として活用することで、群衆量の予測能力を強化する。また、マクロ群衆量は、少人数の行動が変動に大きく影響するミクロ群衆量と比較

して、予測が容易である。そのため、既存の群衆流動予測モデルのように複雑なモデルを利用せずとも、軽量のモデルで高精度に予測することが可能である。

具体的には、このモジュールは、未来の都市全域の群衆量の総量を予測する総量予測ヘッドと、未来の都市全域のマクロ群衆量の分布を予測するマクロ空間分布予測ヘッドから構成される。

総量予測ヘッド：未来の群衆量の総量は、現在の群衆量の総量と強く相関している、すなわち強い自己相関があると考えられる。そこで、総量予測ヘッドでは、直近の群衆量の総量と未来に関する時刻メタ情報を入力に用いる。また、未来の群衆量の総量は、現在の群衆量の総和に対して予測した係数をかけることで求める。具体的には、以下のプロセスで求める。

1. 直近の群衆量の総量、時刻メタ情報に対して、それぞれ全結合層と埋め込み層を適用し、得られた埋め込み表現を結合する。
2. 結合した埋め込み表現に対して、2層の Multi-Layer Perceptron (MLP) を適用することで、直近の群衆量の総和に対する比率を出力する。
3. 出力した比率を、現在の群衆量の総量に対してかけることで、未来の群衆量の総量の予測値を得る。

空間分布ヘッド：未来のマクロ群衆量の分布は、直近の群衆量の分布だけでなく、1日前・1週間前の分布と強く相関していると考えられる。そこで、空間分布予測ヘッドでは、直近・1日前・1週間前の群衆量の観測値と対応するすべてのメタ情報を入力に用いる。ただし、このヘッドでは、あくまでマクロな空間分布の変動に着目するので、各観測値に対してダウンサンプリングを行い、さらに空間軸で平均と標準偏差を求めてインスタンス正規化を行う。このダウンサンプリング・空間インスタンス正規化適用後の観測値に対して、MLP, Multi-Head Attention (MHA), Attention を適用することで、未来の空間分布を求める。具体的には、以下のプロセスで求める。

1. 直近・1日前・1週間前の各観測値と対応するメタ情報の各ペアに対して、観測値とメタ情報のそれぞれに対して埋め込み層を適用し、得られた埋め込み表現を各ペア内で結合する。
2. 結合後の各埋め込み表現に対して、MLP を適用し高次元

空間に射影した後、グリッド軸で MHA を適用することで、他のグリッドの情報を取り込み空間的依存関係を捉えた表現に変換する。また、シンプルな MHA ではなく、明示的にデータの周期的な依存関係を考慮することができる Temporal Query [13] を導入した TQ-MHA を用いる。Temporal Query とは、1 週間などの周期長を持つ学習可能なパラメータであり、Key・Value のデータの曜日・時刻に合わせて周期的にシフトすることで、周期的な空間的依存関係を捉えることを可能にする。

3. ここまでの処理で得られた直近・1 日前・1 週間前に対する埋め込み表現に対して、Attention を用いた加重和による融合を行う。
4. 融合後の埋め込み表現に対して、softmax を適用し、総和が 1 の空間分布に変換する。

最終的に、これら 2 つのヘッドの出力を単純にかけ合わせることで、未来のマクロ群衆量の予測値を得る。

4.2.2 入力変換器

このモジュールの目的は、速度場予測に活用する情報を、後続のモジュールで活用する豊かな初期表現に変換することである。速度場予測に活用する情報とは、直近・1 日前・1 週間前の群衆量・流入量・流出量の観測値データおよび対応するメタ情報、先述のマクロ群衆量予測器の出力である。また、U-Net ベースのモジュールが利用する CNN には、平行移動不変性という特性から、位置情報を明示的に扱うのが不得手であるという欠点がある。そこで、正規化された位置座標情報も、追加情報として用いる。これらの情報を豊かな表現に変換するために、過去の観測値データと対応するメタ情報を変換する **コンテキストエンコーダ**、現在の群衆量と未来のマクロ群衆量を変換する **状態エンコーダ**、座標情報を変換する **座標エンコーダ** を用いる。現在の群衆量は、直近の観測値データにも含まれているが、予測した速度場を適用する初期状態という別の側面も持っているため、2 種類のエンコーダを用いて変換を行う。

コンテキストエンコーダ：直近・1 日前・1 週間前の観測値は、それぞれが時間的に連続したデータ構造を持つ。また、ある時刻の群衆量・流入量・流出量には相互関係が存在する。そこで、まず直近・1 日前・1 週間前の観測値のそれぞれに対して、対応するメタ情報を考慮しつつ群衆量・流入量・流出量をまとめてエンコーディングする。その後、直近・1 日前・1 週間前のエンコーディング結果を、ゲート付き融合を用いて動的に融合する。具体的には、以下のプロセスで求める。

1. 各タイムステップにおいて群衆量・流入量・流出量を結合した後、 1×1 畳み込み層を適用する。
2. 得られた埋め込み表現に対して、メタ情報を埋め込み表現に変換したものを結合し、再度 1×1 畳み込み層を適用する。
3. ここまでの処理で得られた直近・1 日前・1 週間前に対する埋め込み表現に対して、時間軸方向の MHA を適用する。
4. 得られた埋め込み表現に対して、ゲート付き融合を行い、出力する。

状態エンコーダ：現在の群衆量と未来のマクロ群衆量を結合し

た後、 3×3 畳み込みを用いて、周囲の状況を考慮しながら次元拡張し、出力する。

座標エンコーダ：ニューラルネットワークには、低周波成分よりも高周波成分を学習するのが苦手というスペクトルバイアスが存在する。そこで、座標をフーリエ変換して高周波成分を含ませることで、より詳細で微細な位置の違いをモデルが識別できるようにする。具体的には、正規化された位置座標をフーリエ変換した後、 1×1 畳み込みを用いて次元拡張し、出力する。

最終的に、各エンコーダの出力を結合した後、 1×1 畳み込みを用いて融合する。

4.2.3 時空間特徴抽出器

このモジュールの目的は、入力変換器によって得られた豊かな初期表現から時空間的な特徴を抽出し、速度場予測に適した高次元空間へ射影することである。この目的を果たすために、局所的な特徴と大域的な特徴の両者を捉えるのに優れた、階層的な構造を持つ U-Net ベースの Encoder-Decoder を用いる。具体的には、Decoder 部分でマルチスケールの情報を統合して用いる U-Net 3+ [14] を用いる。また、Encoder-Decoder のすべてのブロックには、計算コストを抑えつつ大域的な特徴を考慮できる ConvNeXt V2 [15] を 2 つ積層して用いる。これらの具体的なアーキテクチャは、原論文 [14] [15] を参照されたい。

4.2.4 速度場予測器

このモジュールの目的は、時空間特徴抽出器によって得られた高次元空間の特徴をもとに、複数の速度場と各速度場に対する重みを予測する。具体的には、以下のプロセスで求める。

1. 高次元特徴に、共通の MLP を適用する。
2. 得られた埋め込み表現に対して、速度場と重みごとに異なる MLP を用いて予測を行う。
3. 重みに関する出力に関しては、Entmax を適用し、各グリッドの重みに関して和が 1 になるように正規化する。

4.2.5 補正器

このモジュールの目的は、次節で説明する方法で速度場と重みから求めた群衆量・流入量・流出量の予測値を補正することである。

まず、マクロ群衆量予測器により予測された未来のマクロ群衆量と整合性が取れるように、群衆量の予測値を補正する。これは、マクロ群衆量に対応するエリアで群衆量予測値の和をとり、補正係数を計算しかけることで行われる。その後、群衆量のみ補正した場合、物理的な整合性が取れなくなるため、流入量・流出量に対しても補正を行う。具体的には、圧力ポアソン方程式の近似解を数値解法を用いて求めることで、補正後の群衆量と連続の式を近似的に満たすように、流入量・流出量に対して補正を行う。圧力ポアソン方程式とは、連続の式を満たすための圧力を求める方程式であり、群衆量補正により発生してしまう不整合をグリッド全体に分散する補正が可能となる。

このモジュールで行われる計算には、学習可能なパラメータは一切用いられない。また、学習中に補正を行うと、学習の効果を適切に評価することが難しくなる。そのため、このモジュールは、推論時にのみ用いる。

4.3 学習戦略

提案フレームワークの学習において、データのエンコーディングと速度場予測を同時に学習することは、比較的難しいタスクである。そこで、速度場予測のための条件を生成するためにはじめに行う、マクロ群衆量予測器の学習と同時に、入力変換器内のコンテキストエンコーダも事前に学習する。具体的には、マスク処理した部分を復元するという自己教師あり学習手法で学習する。その後、他のエンコーダや時空間特徴抽出器や速度場予測器と同時に学習する際は、コンテキストエンコーダ部分のみ学習率を下げる。

また、提案フレームワークは、群衆流動のダイナミクスを用いて推論を行うため、学習時にこのダイナミクスを学習することが望ましい。このダイナミクスは、現在から未来だけでなく、過去のタイムステップ間でも常に成立する関係である。そこで本研究では、学習時に未来予測 ($t \rightarrow t+1$) だけでなく過去再現 $t-n \rightarrow t-n+1$ も行うことで、このダイナミクスの学習を強化する。

また、複雑なダイナミクスを表現するために、出力される複数の速度場は多様である必要がある。一方で、予測される速度場は、単に群衆流動の予測精度を向上させるものでなく、物理的に整合性の取れた移動である必要がある。例えば、群衆が突如湧き出たり消失することは許容されない。そこで本研究では、予測した群衆量・流入量・流出量の予測値の正確性を測る損失項だけでなく、速度場の多様性を測る損失項や連続性を測る損失項も用いる。

具体的には、群衆量の再現値・予測値に対して、以下の4つの損失項を計算する。

- **Grid MSE Loss** : 各グリッドごとに、群衆量の再現値・予測値と真値の間で MSE を計算する。
- **Macro L1 Loss** : マクロ群衆量予測器で用いるスケールで、群衆量の再現値・予測値と真値を集約し L1 Loss を計算する。
- **Total L1 Loss** : すべてのグリッドに関して、群衆量の再現値・予測値と真値を集約し L1 Loss を計算する。
- **Grid FFT Loss** : 各グリッドごとに、群衆量の再現値・予測値と真値をフーリエ変換し、周波数領域での振幅スペクトルの L1 Loss を計算する。

単純にグリッド単位のみ損失を計算するのではなく、マルチスケールで計算することで局所的かつ大域的に真値に近づくことを目指す。

また、流入量・流出量の再現値・予測値に対しては、群衆量の損失でも用いた Grid MSE Loss と Grid FFT Loss を用いる。

さらに、予測した速度場やモデルパラメータに対して、以下の6つの損失項が計算される。

- **Continuity Loss** : 現在の群衆量と予測された速度場から計算される発散が、実際の群衆量の時間変化量と一致するかを測る損失項。モデルが連続の式を満たすように誘導する効果を持つ。
- **Smoothness Loss** : 各速度場の空間的な滑らかさを測る

損失項。速度場において、ノイズのような不自然な動きを抑制する効果を持つ。

- **Cycle Consistency Loss** : 現在の群衆量を予測された速度場で移動させた後、逆向きの速度場を適用したとき、元の群衆量を再現できているかを測る損失項。予測した速度場による移動が、時間的に可逆であること、すなわち行って戻ってこられる移動であるように誘導する効果を持つ。ある地点の群衆を消失させ、他の地点に突如出現させるような不連続な移動に対してペナルティを与える。
- **Diversity Loss** : 予測された複数の速度場に関して、ユークリッド空間での距離に基づく分離を最大化するための損失項。異なる速度場間の速度ベクトルの距離を測り、近すぎる場合ペナルティを与える。すべての速度場が似たような速度場に収束することを防ぐ効果を持つ。
- **Repulsion Loss** : 予測された複数の速度場に関して、極座標空間での距離に基づく分離を最大化するための損失項。速度ベクトルを速さと向きに分解し、両方が似ている場合ペナルティを与える。すべての速度場が似たような速度場に収束することを防ぐ効果を持つ。向きが同じでも速さが違うなら異なる挙動とみなすという、Diversity Loss より繊細な区別を可能にする。一方で、2つの速度ベクトルがほぼゼロの場合でも角度が異なれば、完全に異なる挙動として扱われ多様性を確保したとみなされる。すなわち、実用上は意味のない極微小な違いでも多様性があると判断される可能性がある。そのため、Diversity Loss と併用することが望ましい。
- **Orthogonality Loss** : 予測された速度場ではなく、予測する直前のモデルの重み行列を用いて計算する損失項。異なる速度場を出力するための重みベクトル同士が直交するように制約をかける。各速度場が、互いに独立した特徴を学習するように誘導し、表現の冗長性を排除する効果を持つ。

本研究では、以上の損失項を足し合わせて、学習時に最小化する損失関数として扱う。しかし、これらの損失項は、群衆量の過去再現・未来予測に対して 2×4 個、流入量・流出量の過去再現・未来予測に対して $2 \times 2 \times 2$ 個、予測した速度場やモデルパラメータに対して6個の、合計22個存在する。そのため、これらの損失項に対して、それぞれ重みをハイパーパラメータチューニングすることはコストがかかる。しかし、適切な重みが設計できなければ、値が大きい損失項が学習を支配し、他の損失項に関する学習が疎かになる可能性がある。

そこで本研究では、群衆量・流入量・流出量に関する16個の損失項に対して Gradient Normalization (GradNorm) [16] を適用する。GradNorm は、各損失項の勾配の大きさを監視し、学習速度が均等になるように重みを動的に調整する手法であり、すべての損失項に対する学習をバランスよく行うことを促す。一方で、予測した速度場やモデルパラメータに対する損失項は、GradNorm の適用外とする。これらの損失項は、速度場やモデルパラメータが妥当であるための正則化や制約条件であるためである。

表 1 ベースライン手法と提案手法の定量的比較

Model	Metric	MSE	RMSE	MAE	WAPE	不整合率
DeepMeshCity	群衆量	29.63	5.44	3.13	4.72	
	流入量	20.51	4.53	2.53	22.33	13.24
	流出量	12.13	3.48	1.92	16.96	
提案フレームワーク	群衆量	39.14	6.09	3.40	5.16	
	流入量	41.90	6.29	3.45	30.62	5.52
	流出量	32.79	5.47	2.77	24.63	
提案フレームワーク w/o 補正器	群衆量	69.22	7.98	3.82	5.81	
	流入量	36.83	5.87	3.02	26.84	10.32
	流出量	33.76	5.61	2.63	23.34	

GradNorm を導入することで、ハイパーパラメータチューニングのコストを減らしつつ、学習速度のバランスをとることが可能となる。しかし、依然として、群衆量・流入量・流出量予測という複数タスクの学習を行う状況下では、タスク間の勾配衝突という問題が残される。たとえ勾配の大きさが揃っていたとしても、タスクごとに勾配が逆方向を見ている場合、単純な加重和で更新すると、互いに干渉し学習の停滞・一方のタスクの悪化を引き起こしうる。

そこで本研究では、モデルの汎化性能を向上させるための最適化アルゴリズム Sharpness-Aware Minimization (SAM) [17], 全てのタスクを同時に改善できる共通の降下方向をみつける Minimum Norm Solver (MinNormSolver) [18] を導入する。これにより、タスク間の不均衡・勾配衝突を回避しつつ、汎化性能が高くパレート最適解を目指すことができる。

5 実 験

本節では、提案モデルを評価するための実験について説明する。具体的には、実験に使用するデータセットの概要と前処理について説明する。その後、実験設定について説明した後、実験結果および考察について述べる。

5.1 データセット

本研究では、LINE ヤフー株式会社によって公開されている群衆流動データ [4] を用いる。記録されているデータの種類は、群衆量・流入量・流出量の 3 種である。また、観測期間は 2017 年 4 月 1 日から 7 月 9 日の 100 日間であり、30 分間隔で記録されている。一方、観測範囲は東京 23 区を中心とした北緯 35.50 から 35.82 度、東経 139.50 から 139.90 度の範囲であり、約 450 m 四方の 80 × 80 個のグリッドに分割されている。

物理法則を満たすための前処理 このデータセットは、群衆量の増減が流入量・流出量の差分に一致するという物理法則を常に満たすというわけではない。これは、観測ノイズやプライバシー保護のための処理に起因すると考えられる。このノイズによる物理法則を満たさないデータは、第 2.2 節で述べたように物理法則を組み込むモデルにとって大きな問題となる。

そこで、本研究では、データを物理法則を満たすように前処理を行い、学習・評価を行う。前処理には、既存研究 [10] [11]

と同様に、最適化ベースの射影を用いる。具体的には、物理法則を制約条件とした凸二次計画問題を 1 日単位で解くことで、物理法則を満たすデータを取得する：

学習・検証・テストデータの分割： 既存研究 [4] [5] に倣い、64 : 16 : 20 の比率で時系列順に分割し、それぞれ学習・検証・テストデータとする。さらに、本研究では物理法則を満たさないデータを削除するために、1 日の最初と最後のデータを削除する。

6 評価指標・比較手法

群衆量・流入量・流出量予測に関してはグリッド単位の平均二乗誤差 (MSE)・二乗平均平方根誤差 (RMSE)・平均絶対誤差 (MAE)・加重絶対パーセント誤差 (WAPE) を用いる。WAPE は、真値が 0 の場合計算できない平均絶対パーセント誤差 (MAPE) に対して、WAPE は以下の式で計算されるため真値が 0 の部分が存在しても計算可能である：

$$WAPE = \frac{\sum |Pred - True|}{\sum |True|} \times 100. \quad (6)$$

さらに、物理的な整合性を、群衆の移動量に対してどの程度不整合かパーセント評価するための指標を導入することで測る。具体的には、質量保存則を表す等式 (3) の左辺と右辺の残差の絶対値を用いて、以下のように定義する：

$$\text{不整合率} = \frac{\sum |Residual|}{\sum |Inflow_{t+1} + Outflow_{t+1}|} \times 100. \quad (7)$$

また、比較手法として既存の群衆流動予測モデル DeepMeshCity [5] を用いる。このモデルは、実験に使用するデータにおいて、現在最も良い性能を達成している群衆流動予測モデルである。既存研究に倣い、群衆量予測と流入量・流出量予測に関して、個別に学習・推論を行う。さらに、アブレーション実験として、提案手法から補正器を抜いたものを用いる。

7 実験結果・考察

提案手法および比較手法の評価結果を表 1 に示す。

まず、提案手法と DeepMeshCity の比較を行う。WAPE に着目すると、提案手法は群衆量の予測において DeepMeshCity に近い予測精度を実現している一方で、流入量・流出量の予測

精度は大きく劣っている。しかし、物理的不整合率に着目すると、提案手法が DeepMeshCity に比べて優れていることがわかる。このことは、提案手法が物理法則を適切に考慮することができ、物理的な整合性のある予測に関して大きな優位性があることを示している。

次に、提案手法と補正層を除いたアブレーション実験の比較を行う。WAPE に着目すると、補正層を導入することで、流入量・流出量の予測精度は少し劣化するが、群衆量の予測精度は改善することがわかる。さらに、物理的不整合率にも明らかな改善がみられる。このことは、補正器を適用する必要性を示している。補正器には学習可能なパラメータがなく、推論時に適用するのみでよいため、補正器はコスト対効果が非常に優れたモジュールだと考えられる。

8 結 論

本研究では、既存の群衆流動予測モデルが物理的に不整合な予測をしまう問題に対処するために、物理法則を組み込んだ群衆流動予測フレームワークの提案を行った。具体的には、群衆流動に内在する物理法則であるが既存の予測モデルでは明示的に考慮されてこなかった、あるエリアの群衆量の増減は流入量と流出量の差分に一致するという関係性に着目し、これを考慮し複数の速度場を通じて予測を行うフレームワークの提案を行った。提案フレームワークを、実世界群衆流動データで学習・評価した結果、物理的整合性を保ちつつ群衆量を高精度で予測できることが明らかになった。一方で、流入量・流出量に関しては改善の余地が残る結果となった。現在のモデル構造では、流入量・流出量に関しては入力データとしてのみ活用されているため、1週間前の観測値などを予測の際に直接参照するショートカット接続を導入などにより、この問題を緩和することが、今後の研究課題である。

謝 辞

本研究の一部は JSPS 科研費 JP24K03046 の助成を受けたものです。

文 献

- [1] Renhe Jiang, Du Yin, Zhaonan Wang, Yizhuo Wang, Jiewen Deng, Hangchen Liu, Zekun Cai, Jinliang Deng, Xuan Song, and Ryosuke Shibasaki. DL-traffic: Survey and benchmark of deep learning models for urban traffic prediction. In *Proceedings of the 30th ACM international conference on information & knowledge management*, pp. 4515–4525, 2021.
- [2] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, Vol. 28, , 2015.
- [3] Yipeng Liu, Haifeng Zheng, Xinxin Feng, and Zhonghui Chen. Short-term traffic flow prediction with conv-lstm. In *2017 9th international conference on wireless communications and signal processing (WCSP)*, pp. 1–6. IEEE, 2017.
- [4] Renhe Jiang, Zekun Cai, Zhaonan Wang, Chuang Yang,

- Zipei Fan, Qunjun Chen, Kota Tsubouchi, Xuan Song, and Ryosuke Shibasaki. Deepcrowd: A deep model for large-scale citywide crowd density and flow prediction. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 35, No. 1, pp. 276–290, 2021.
- [5] Chi Zhang, Linhao Cai, Meng Chen, Xiucheng Li, and Gao Cong. Deepmeshcity: A deep learning model for urban grid prediction. *ACM Transactions on Knowledge Discovery from Data*, Vol. 18, No. 6, pp. 1–26, 2024.
- [6] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, Vol. 3, No. 6, pp. 422–440, 2021.
- [7] Zhongkai Hao, Songming Liu, Yichi Zhang, Chengyang Ying, Yao Feng, Hang Su, and Jun Zhu. Physics-informed machine learning: A survey on problems, methods and applications. *arXiv preprint arXiv:2211.08064*, 2022.
- [8] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, Vol. 378, pp. 686–707, 2019.
- [9] Shuai Han, Lukas Stelz, Horst Stoecker, Lingxiao Wang, and Kai Zhou. Approaching epidemiological dynamics of covid-19 with physics-informed neural networks. *Journal of the Franklin Institute*, Vol. 361, No. 6, p. 106671, 2024.
- [10] Liang Zhang, Gang Wang, and Georgios B Giannakis. Real-time power system state estimation and forecasting via deep unrolled neural networks. *IEEE Transactions on Signal Processing*, Vol. 67, No. 15, pp. 4069–4077, 2019.
- [11] Jared L Callahan, James V Koch, Bingni W Brunton, J Nathan Kutz, and Steven L Brunton. Learning dominant physical processes with data-driven balance models. *Nature communications*, Vol. 12, No. 1, p. 1016, 2021.
- [12] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.
- [13] Shengsheng Lin, Haojun Chen, Haijie Wu, Chunyun Qiu, and Weiwei Lin. Temporal query network for efficient multivariate time series forecasting. *arXiv preprint arXiv:2505.12917*, 2025.
- [14] Huimin Huang, Lanfen Lin, Ruofeng Tong, Hongjie Hu, Qiaowei Zhang, Yutaro Iwamoto, Xianhua Han, Yen-Wei Chen, and Jian Wu. Unet 3+: A full-scale connected unet for medical image segmentation. In *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 1055–1059. Ieee, 2020.
- [15] Sanghyun Woo, Shoubhik Debnath, Ronghang Hu, Xinlei Chen, Zhuang Liu, In So Kweon, and Saining Xie. Convnext v2: Co-designing and scaling convnets with masked autoencoders. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 16133–16142, 2023.
- [16] Zhao Chen, Vijay Badrinarayanan, Chen-Yu Lee, and Andrew Rabinovich. Gradnorm: Gradient normalization for adaptive loss balancing in deep multitask networks. In *International conference on machine learning*, pp. 794–803. PMLR, 2018.
- [17] Pierre Foret, Ariel Kleiner, Hossein Mobahi, and Behnam Neyshabur. Sharpness-aware minimization for efficiently improving generalization. *arXiv preprint arXiv:2010.01412*, 2020.
- [18] Ozan Sener and Vladlen Koltun. Multi-task learning as multi-objective optimization. *Advances in neural information processing systems*, Vol. 31, , 2018.