ABN と Kolmogorov-Smirnov 検定によるデータのドリフト検知

EP19080 新田常顧　　指導教授：山下隆義

1. はじめに

データのドリフトは，運用時のデータ分布がモデル学習時のデータ分布から時間とともに変化することであり，その検知は重要な課題である．ドリフトは，カメラのレンズの経年劣化により発生するノイズや，カメラの交換によって発生する位置ズレや反転などにより生じる．このようなドリフトは，機械学習モデルの性能低下の原因となる．そのため，データ分布の変化を監視することが重要となる．本研究ではモデルの判断根拠の可視化が可能である Attention Branch Network [1] と 2 標本検定によるドリフト検知法を提案する．

2. 関連手法

ドリフト検知は，モデル学習時のデータ分布と運用時のデータ分布を比較することで行われる．Rabanser らの手法 [2] は，それぞれのデータに対して次元削減を行い，2 標本検定により分布を比較してドリフトを検知する．学習済みモデルが出力するクラス確率分布に対して，2 標本検定の一種である Kolmogorov-Smirnov (KS) 検定を用いる検知手法が高精度であることを示した．

3. 提案手法

本研究では，ABN を用いたドリフト検知法を提案する．ABN では，Attention branch から得られるアテンションマップを Attention 機構に入力し，特定領域の特徴を強調して推論を行う．そのため，ドリフトによるアテンションマップの変化によって，入力画像の変化に対する特徴が強調され，ドリフトの検知精度が向上すると考える．例えば，図 3 に示すようにドリフトした画像に対するアテンションマップは，ドリフトなしと比較して変化する．このことから，アテンションマップの違いを考慮することで，ドリフト検知の高精度化が期待できる．ABN を用いたドリフト検知の流れを図 1 に示す．また，提案手法は以下のデータセットの前処理と Step1 から Step3 の流れでドリフト検知を行う．

データセットの前処理

データセットを学習用データ，検証用データ，テスト用データの 3 つに分割する．検証用データは学習用データと同じデータ分布であると仮定し，ドリフトなしのデータとする．テスト用データはモデル運用時のデータ分布と仮定し，ドリフトありのデータとする．

Step1. ABN の学習

学習用データを用いて ABN を学習する．

Step2. クラス確率分布とアテンションマップの算出

ドリフトなしのデータとドリフトありのデータから，指定したサンプル数を取得し，学習済み ABN に入力する．それぞれのデータ群に対して Attention branch の出力するクラス確率分布を (D^{Ab}, D^{Ab}_{drift})，アテンションマップを (D^{Am}, D^{Am}_{drift})，Perception branch の出力するクラス確率分布を (D^{Pb}, D^{Pb}_{drift}) とし，保存する．

Step3. ドリフト検知

Step2 で求めた各分布に対して，式 (1) に示す KS 検定を行い，p 値 (p_{Ab}, p_{Am}, p_{Pb}) を求める．

$$\begin{aligned} p_{Ab} &= \mathrm{KS}(D^{Ab}, D^{Ab}_{drift}) \\ p_{Am} &= \mathrm{KS}(D^{Am}, D^{Am}_{drift}) \\ p_{Pb} &= \mathrm{KS}(D^{Pb}, D^{Pb}_{drift}) \end{aligned} \tag{1}$$

p 値は，2 つの分布の母集団が同じであるという仮定の下で，KS から得られる検定統計量がそれ以上の値となる確率である．次に，式 (2) に示すように p 値の最小値を p_{min} とする．

$$p_{min} = \min(p_{Ab}, p_{Am}, p_{Pb}) \tag{2}$$

最後に，$p_{min} < \alpha$ であればドリフトありと判定する．ここで，α は従来手法と同値の閾値である．

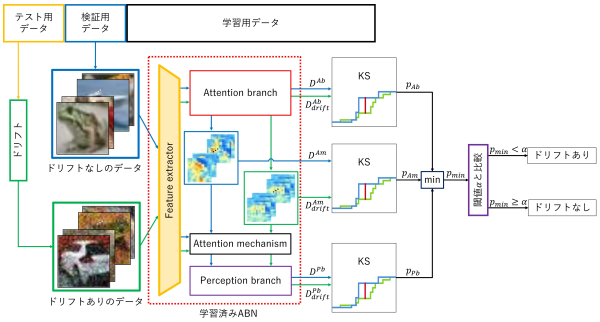


図 1: 提案手法によるドリフト検知の流れ

4. 評価実験

提案手法の有効性を調査するために，従来手法とドリフト検知精度の比較を行う．

4.1. 実験概要

本実験では，MNIST データセットと CIFAR-10 データセットを用いてドリフト検知を行う．各データセットの分割方法やデータの扱いは Rabanser らの手法 [2] に従う．ドリフトの再現には，ガウシアンノイズ，幾何変換，データの不均衡化を用いる．また，ドリフト検知のために必要とするサンプル数を評価するために，ドリフト検知に用いるサンプル数を 6 種類用意する．ここで，評価指標には各種ドリフトの再現手法での平均検知率を用いる．

4.2. 実験結果

Rabanser らの手法と提案手法の実験結果を図 2 に示す．図 2 より，ドリフト検知に用いるサンプル数に関わらず平均検知率が向上している．また，サンプル数が多いほど検知率が向上している．画像にガウシアンノイズを付与した際のアテンションマップを図 3 に示す．図 3 より，ガウシアンノイズをによるドリフトの際は注視領域が狭まっていることが分かる．ABN はアテンションマップを用いて推論を行うため，これらの変化を得ることでドリフトの検知精度が向上していると考えられる．

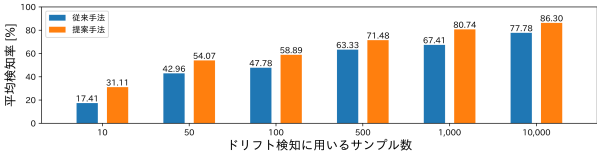
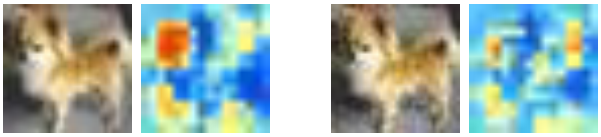


図 2: 提案手法と従来手法の平均検知率の比較



ドリフトなし　　ガウシアンノイズ

図 3: アテンションマップの比較

5. おわりに

本研究では，ABN を用いてドリフトを検知する手法を提案し，ドリフト検知精度が向上することを確認した．また，ドリフトに対するアテンションマップの変化を分析した．今後は，その他のドリフト検知手法との比較や，アテンションマップを用いたドリフト検知，ドリフトの再現手法を追加し，実験を行うことを検討する．

参考文献

[1] H.Fukui, *et al.*, “Attention Branch Network: Learning of Attention Mechanism for Visual Explanation”, CVPR, pp.10705-10714, 2019.

[2] Rabanser, *et al.*, “Failing Loudly: An Empirical Study of Methods for Detecting Dataset Shift”, NeurIPS, 2019.