

1. はじめに

ネットワークの枝刈りは、重要度の低いニューロンを削除することで、ネットワークを軽量化する手法である。従来の枝刈り手法の多くは、静止画像を対象としており、時系列データに最適な手法が提案されていない。そこで、本研究では時系列データを対象とした、移動エントロピーによる枝刈り手法を提案する。これにより、再帰型ニューラルネットワーク (RNN) の分類精度を維持しつつニューロンを削減できることとする。

2. 従来手法

Liu ら [1] は、エントロピーや相互情報量を算出し、認識精度への寄与率 (重要度) が低いと考えられるニューロンを削減する手法を提案している。エントロピーはニューロンの出力から算出し、相互情報量はニューロンの出力と正解ラベルから算出する。尚この手法では、静止画像の識別を対象としており、動画のような時系列データを対象としていない。

3. 提案手法

本研究では、時系列データを対象とし、移動エントロピー [2] を用いた RNN の枝刈り手法を提案する。移動エントロピーは情報量的一种であり、情報の流れを抽出することが可能である。提案手法は、図 1 に示すように、中間層の時刻 t の出力を x_t 、時刻 $t+1$ の出力を x_{t+1} 、その学習データに対応する時刻 t の正解ラベルを y_t とすると、移動エントロピーは式 (1) により算出できる。

$$T_{Y \rightarrow X} = \sum_{x_{t+1}, x_t \in X} \sum_{y_t \in Y} p(x_{t+1}, x_t, y_t) \log \frac{p(x_{t+1} | x_t, y_t)}{p(x_{t+1} | x_t)} \quad (1)$$

$Y \rightarrow X$ は、 Y から X への因果的な影響の情報理論的な定量化を示す。これにより、エントロピーや相互情報量では得られなかった時系列性を考慮した重要度を得ることができる。この移動エントロピーの値の大きさに基づき削減するニューロンを決定することで、RNN の軽量化を図る。

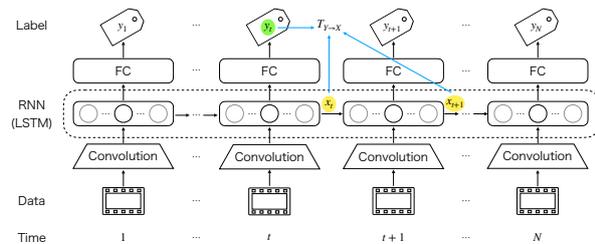


図 1: ネットワークの構造と移動エントロピーの算出

ネットワークは図 1 に示すように、畳み込み層 (Convolution), RNN 層, 全結合層 (FC) からなるネットワークである。畳み込み層のチャンネル数は 16 である。RNN 層には Long Short Term Memory (LSTM) を 128 ユニット用いる。全結合層のユニット数は対象とするタスクまたはデータセットのクラス数と同一である。

4. 評価実験

移動エントロピーによる枝刈り手法の有効性を示すために、相互情報量と移動エントロピーで枝刈りした場合の精度比較を行う。

4.1. UIUC Video Dataset

評価データは、UIUC Video Dataset を用いる。このデータセットは、静止画像の UIUC Dataset のテキストチャ画像を使用し、各画像に対して上下左右の移動、拡大縮小、左右回転の動きを加えた動画から構成される。訓練データは 32,000 本、評価データは 8,000 本、フレームサイズは 32×32 pixels、フレーム数は 100 である。1 つの動画中に異なる動きが含まれており、1 つの動作のフレーム数は 50 である。

4.2. 実験概要

まず、ネットワークを UIUC Video Dataset を使用して動きのクラス識別の学習を行う。その後、そのネットワークに対して 2 種類のニューロンの削減方法を行う。一つ目は、RNN の中間層の各ニューロンに対して、特定のニューロンを削減した際の精度を調査する。二つ目は、重要度の低いニューロンを削減した際の精度を調査する。

4.3. 実験結果

図 2 に RNN 層 (ニューロン数:128) の各ニューロンの相互情報量を x 軸、移動エントロピーを y 軸とし、特定のニューロンを削減した際の精度のカラマップを示す。ここで、ニューロン n_1 と n_2 に着目すると、これらのニューロンは他のニューロンを削減した際と比べ、精度が大幅に低下することがわかる。 n_1 は相互情報量が低く、 n_2 は高い。一方で、移動エントロピーは共に低い。これより、移動エントロピーが低いニューロンが重要であると考えられる。したがって、相互情報量よりも移動エントロピーの方が認識結果への寄与率が高いといえる。

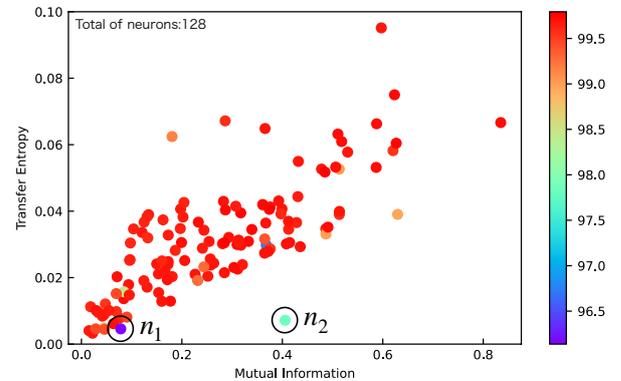


図 2: 各ニューロンの相互情報量と移動エントロピー

図 3 に RNN 層 (ニューロン数:128) のニューロンの内、相互情報量と移動エントロピーの値の高いニューロンを削減した場合 (Low) と、低いニューロンを削除した場合 (High) の認識精度の比較を示す。共に High と比べ、Low が精度が安定していることがわかる。これにより、情報量が低いほど重要なニューロンであるといえる。これは正解ラベルとの曖昧さが少ないためだと考えられる。また、相互情報量よりも移動エントロピーを用いた方が高い認識精度を維持している。これにより、相互情報量では得られなかった重要度を移動エントロピーが獲得しているといえる。

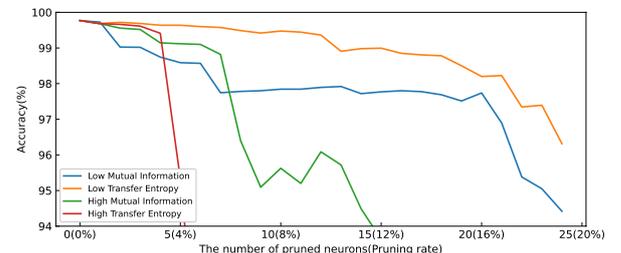


図 3: RNN の枝刈りによる精度の比較

5. おわりに

本研究では移動エントロピーを用いた RNN の枝刈り手法を提案した。今後の予定として、異なる指標を用いた RNN 層のニューロンの影響の分析などが挙げられる。

参考文献

[1] K. Liu, *et al.*, “Understanding individual neuron importance using information theory.” arXiv preprint arXiv:1804.06679, 2018.
 [2] T. Schreiber, “Measuring information transfer”, *Phys. Rev. Lett.*, Vol.85, No.2, pp. 461–464, 2020.