

深層学習を用いた動画像からの危険認知手法のための基礎的研究

4 月 14 日 (火)

小松 大起

1 PredNet

PredNet は Deep Recurrent Convolutional Neural Network の 1 種で 神経科学の概念である Predictive Coding を組み込んで作られたモデルである。

1.1 Predictive Coding

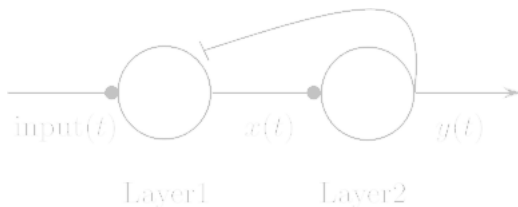


Fig. 1: 2 つのニューロンの模式図

2 種類のニューロンが存在するとし、それを図 1 で示す。図の上では 2 つのニューロンであるが、Layer1 と Layer2 という 2 層について、それぞれを代表するニューロンを 1 つずつ描いている。まず、Layer1 のニューロンは下層から $input(t)$ を受け取る。これは入力刺激と呼ばれる。次に Layer1 のニューロンは $x(t)$ を出力する。これは誤差信号または残差と呼ばれ、以下の式が成り立つ。

$$x(t) = input(t) - y(t) \quad (1)$$

また、 $y(t)$ は Layer2 のニューロンからの出力で、Layer1 のニューロンへの入力を予測する。ここで

$$dy/dt = x(t) = input(t) - y(t) \quad (2)$$

であり、 $dy/dt = 0$ すなわち $y(t) = input(t)$ となるように学習が進行するモデルである。

1.2 PredNet の層構造

PredNet の各層には 4 つの素子が存在しておりそれぞれ、Target, Representation, Prediction, Error と呼ぶ。Target は下層からの出力である誤差信号をエンコード、符号化する。Representation は Recurrent unit で、上層からの出力、側方からの誤差信号、1 ステップ前の自分の出力を受け取る。Representation unit は Target の予測をする Prediction unit に投射し、入力の予測が出力される。Error は Prediction と Target の誤差であり、Error は上層に送られる。この Error が小さくなるように学習を進めていく。また、層構造の例を図 2 に示す。

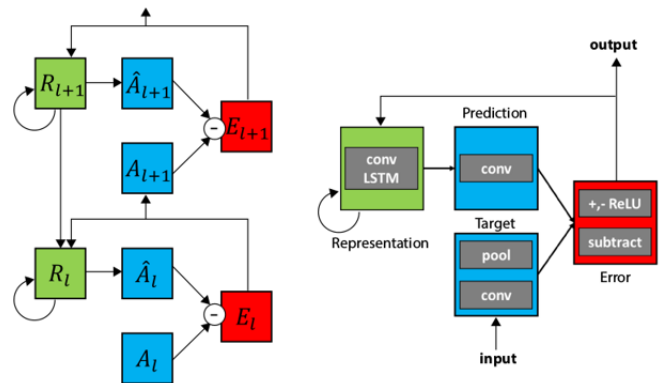


Fig. 2: PredNet の層構造の例

2 先週までの作業

- 全体ゼミの予習
- PredNet について

3 今週の作業

- PredNet について調べる Deep Predictive Coding Networks for Video Prediction and Unsupervised Learning という論文で提案された手法
- 予測に用いる動画像探し kitti というサイトにある動画像を使えるのではないかな

4 来週以降の作業

- もう少し動画像を集める, 日本の道の動画像も欲しい
- プログラムを読む