

算数からはじめる 画像ディープ・ラーニング

ご購入はこちら

牧野 浩二, 西崎 博光

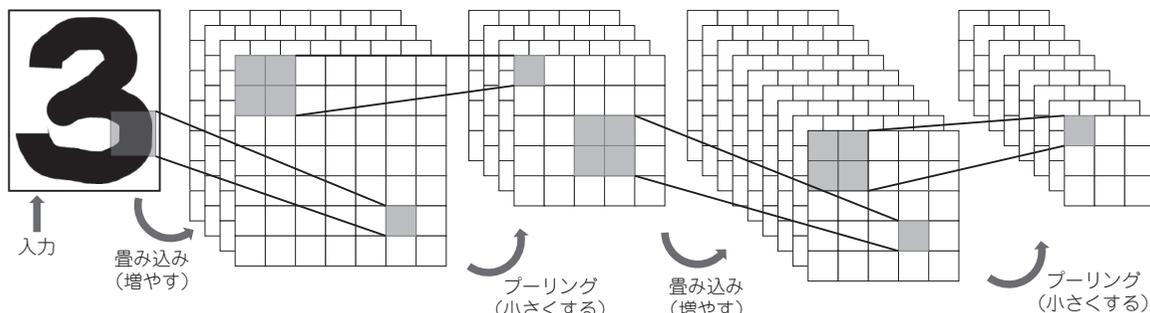


図1 畳み込みニューラル・ネットワークで肝となるフィルタ処理は2つ

いきなり数式の話をしてしまします。3大ディープ・ラーニングの1つである「畳み込みニューラル・ネットワーク」を使いこなそうとすると、後述する、

$$y = \frac{x + 2p - f}{s} + 1$$

の各パラメータが整数となるように設定しなければ、計算ができません。

ニューラル・ネットワークの層が深くなると、これがいつでも整数になるような p , f , s を選ぶのは難しくなります。いろいろな本にもそのように書かれています。なので、この方法をマスターする必要がありました。しかし、Chainerは整数にならなくてもうまくやってくれます。

畳み込みニューラル・ネットワークは、絶対に数式まで理解しないと使えないアルゴリズムでしたが、ChainerやTensorFlowのおかげで、大まかな内容を理解しておくだけで済むようになりました。ただし、これから説明するキーワードを知らなければ、設定はできません。

畳み込みニューラル・ネットワークは、図1のように、リカレント・ニューラル・ネットワークやオート・エンコーダとは違った計算をします。

ここで重要となるのが「フィルタ」というものです。この「フィルタ」を用いて画像の処理と画像の縮小を行います。この計算方法を知っておかないと、畳み込みニューラル・ネットワークの設定がうまくできませ

ん。そこで、フィルタを用いた「画像の処理とは何か」と「画像の縮小とは何か」を1つ1つ説明していきます。

「畳み込み」とは

■ 基本中の基本：フィルタ処理

● まずは左上の3×3を対象に

畳み込みの仕組みは図と対応させながら理解すれば、全く難しくありません。図2(a)のような6×6画素の画像を例として用います。画像には0以上の数が入っているものとします。畳み込みにはフィルタというものがようになります。これは図2(b)の例では3×3のフィルタを示しています。このフィルタに使う数字はマイナスもOKとなっています。

3	4	1	2	7	2
2	4	2	3	8	6
7	3	4	2	5	6
2	5	6	7	1	3
4	5	2	1	7	2
1	8	3	9	6	3

(a) 入力画像は
6×6画素

1	-2	1
-2	1	-1
2	2	-1

(b) 3×3の
フィルタ

図2 6×6画素の画像に3×3のフィルタを適用