

## 卷積神經網路原理及其C++/Opencv實作(5)—參數更新

原創 sdf20201029 萌萌噠程序猴 2021-03-26 16:38

上一篇文章我們講了誤差訊息的反向傳播過程，核心思想在於複合函數的鍊式求導法則：

卷積神經網路原理及其C++/Opencv實作(4)—誤反向傳播法

本文我們主要講怎麼使用誤反向傳播過程中的局部梯度資訊來更新神經網路的參數。5層網路需要更新調節的參數主要包括：

1. C1層的6個5\*5卷積核，以及6個偏壓值。
2. C3層的6\*12個5\*5卷積核，以及12個偏移值。
3. O5層的192\*10個權重值，以及10個偏移值。

首先我們來回顧一下神經網路的正向傳播過程，以下我們只列出公式，具體在前文已經講過：

卷積神經網路原理及其C++/Opencv實作(3)

### 1. C1層

C1的前向傳播公式如下，其中 $0 \leq i < 6$ 。

$$y_{C1}^i = I * k_{C1}^i + b_{C1}^i$$
$$Y_{C1}^i = f(y_{C1}^i)$$

### 2. S2層

S2的前向傳播公式如下。

$$Y_{S2}^i = pool(Y_{C1}^i), 0 \leq i < 6$$

### 3. C3層

C3的前向傳播公式如下。

#### 4. S4層

S4的前向傳播公式如下。

#### 5. O5層

O5的前向傳播公式如下。

#### 6. 交叉熵誤差函數

交叉熵誤差函數如下，其中 $t$ 為標籤：

要使用梯度下降法優化參數，關鍵在於求參數關於交叉熵誤差函數的偏導數。

首先，我們來求O5層的權重和偏壓的偏導數，如下式，其中 $0 \leq i < 10$ ， $0 \leq j < 192$ 。

看到上式中 $E$ 關於 $y$ 的偏導數，也許有人不理解了，這裡我們在上篇文章中已經詳細推導過（Softmax函數求導），讀者可以點開本文開頭的超連結參考上篇文章哦~

求偏導數之後，就可以更新參數了，其中 $\alpha$ 為學習率，需要根據經驗設定一個合適的初始值，通常隨著梯度下降法的迭代而逐漸減少。

其次，我們來求C3層參數的偏導數，如下式，其中 $Y$ 、 $y$ 、 $k$ 、 $d$ 都是二維矩陣。'\*'表示兩個矩陣的Valid模式捲積，"."表示兩個矩陣對應位置的值相乘，DerivativeRelu表示Relu函數的導數。

上式中， $0 \leq i < 12$ ， $0 \leq j < 6$ ，結果輸出 $(12-8+1) \times (12-8+1) = 5 \times 5$ 的矩陣，其中 $d_{S4}$ 為S4層的局部梯度（見上篇文章）。這裡可能有人還會有疑問， $y$ 關於 $k$ 的偏導數為什麼放到卷積符號的左邊了呢？以下我們舉個簡單例子來說明卷積求導的公式（具體推導過程後續再研究）。

例如我們有矩陣 $A$ 、 $B$ 、 $C$ 、 $X$ ，以 $X$ 為卷積核對 $A$ 進行卷積得到 $B$ ， $C$ 是 $B$ 經由函數 $f$ 運算之後得到的矩陣，也即：

$$B = A * X$$

$$C = f(B)$$

那就求 $C$ 關於 $X$ 的偏導數，按下式計算，可以看到 $B$ 關於 $X$ 的偏導數是 $A$ ， $A$ 本來在卷積符號的左側，求導時還是在左側。

若求 $C$ 關於 $A$ 的偏導數，按下式計算，可以看到 $B$ 關於 $A$ 的偏導數是 $X$ ， $X$ 本來在卷積符號的右側，求導時還是在右側，不過求導時需要對 $X$ 進行順時針 $180^\circ$ 旋轉。

以上是求 $E$ 關於卷積核 $k$ 的偏導數，接下來求 $E$ 關於偏壓 $b$ 的偏導數。我們知道，卷積結果是一個二維矩陣，該矩陣加上偏壓的操作，相當於矩陣中每個值都加上同一個偏移值。如下圖所示：

在這裡，我們針對 $y_{C3}$ 矩陣的每一個值 $y_{C3}(x,y)$ 來計算。首先我們知道矩陣 $y_{C3}$ 的偏導數為：

則對於 $0 \leq x < 8$ ， $0 \leq y < 8$ 的每一個 $y_{C3}(x,y)$ 來說，其偏導數為：

由上述可知，偏壓 $b$ 與卷積結果 $y_{C3}$ 矩陣中每一個值 $y_{C3}(x,y)$ 都有關，從而得出偏壓 $b$ 的偏導數如下，其中 $0 \leq i < 12$ 。

求偏導數之後，更新參數如下：

最後，我們來求C1層參數的偏導數。與C3層的計算方法類似，如下式，其中 $I$ 、 $Y$ 、 $y$ 、 $k$ 、 $d$ 都是二維矩陣。 $*$ 表示兩個矩陣的Valid模式捲積， $.$ 表示兩個矩陣對應位置的值相乘， $\text{DerivativeRelu}$ 表示Relu函數的導數。

上式中， $0 \leq i < 6$ ，結果輸出 $(28-24+1)*(28-24+1)=5*5$ 的矩陣，其中 $d_{s2}$ 為S2層的局部梯度（見上篇文章）。

接下來求E關於偏壓b的偏導數，與C3層的計算過程類似，其中 $0 \leq i < 6$ ：

求偏導數之後，更新參數如下：

好了，本文我們就講到這裡，在接下來的文章中，我們會詳細講怎麼使用C++和Opencv來實現這5層網路。

歡迎掃碼追蹤以下微信公眾號，接下來會不定時更新更加精彩的內容噢～

人工智慧 27    深度學習 26    機器學習 33

人工智慧 目錄

上一 篇

卷積神經網路原理及其C++/Opencv實作(4)  
一誤反向傳播法

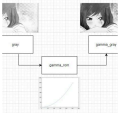
下一 篇

卷積神經網路原理及其C++/Opencv實作(6)  
一前向傳播程式碼實現

閱讀原文

喜歡此內容的人還喜歡

數位影像處理之gamma矯正  
FPGA開源工作室



NJ系列電子凸輪應用分享  
Karl工控



混凝土模板荷載與壓力計算  
忒修斯破船



