

Rot-CNN:一種抗旋轉的卷積神經網路

参展學生: 楊佳誠

指導教授: 陳弘軒 老師

專題類型:■人工智慧、資料科學與多媒體研究

> 專題簡介

- CNN 通常需要透過 Data Augmentation 的方式來辨識經旋轉的圖像
- 我們試圖設計一種方法,能在不進行Data Augmentation的情況下,也能讓模型對稍加旋轉的資料進行正確判別
- 理想情況下,該模型不僅能正確判別旋轉後的圖像,另外還能縮短訓練時間,大幅減少所需訓練的參數量

> 專題成果

A. 資料集說明

使用MNIST、FashionMNIST、CIFAR-10三個經典的資料集進行實驗。三者資料集都是進行十個類別的分類問題。

- MNIST、FashionMNIST: 28*28的灰階圖片組成
- CIFAR-10:32*32大小的彩色圖片組成

資料集又分為下列三種:

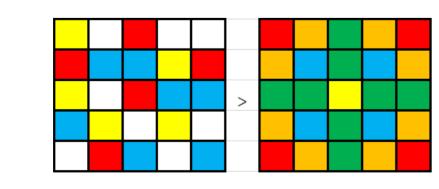
Test:為各個資料集的預設測試資料集
rot90x:將Test資料集旋轉90度的倍數
ranrot:將Test資料集旋轉隨機角度

B. 實驗設定

本實驗透過下列三種模型進行分析比較:

- CNN1:兩層Convolution層, filter大小分別為3*3、5*5, 加上兩層Affine層
- CNN2: 一層Convolution層, filter大小分別為28*28或 32*32, 加上一層Affine層
- Rot-CNN:一層Convolution層, filter大小分別為28*28或32*32,取決於輸入圖片大小,加上一層Affine層。其中此filter在初始化以及倒傳遞後,根據函數flip_filter來調整形狀,讓filter處於對稱。

def flip_filter(FN):
 ANS = torch.rot90(FN, 1, [2, 3])+ torch.rot90(FN, 2, [2, 3])+torch.rot90(FN, 3, [2, 3])+ FN
 x = torch.flip(FN,[3])
 ANS = ANS + torch.rot90(x, 1, [2, 3])+ torch.rot90(x, 2, [2, 3])+torch.rot90(x, 3, [2, 3])+ x



Filter對稱示意圖

C. 評量指標

ANS = ANS/8

return ANS

觀察模型的Accuracy、Average loss、總共所需學習的參數量以及每個epoch所需運行的時間做為比較項目

D. 實驗結果

Accuracy與Average loss:

			MNIST		F	ashionMNIS	T		CIFAR-10	
		Test	rot90x	ranrot	Test	rot90	ranrot	Test	rot90	ranrot
CNN1 (data augmentation)	Ave_loss	0.1266	0.1159	0.1330	0.5251	0.5087	0.5251	1.3254	1.3319	1.3282
	Accuracy	96%	96%	96%	81%	82%	81%	52%	53%	52%
CNN1 (without data augmentation)	Ave_loss	0.0223	6.9584	4.3955	0.2297	8.4026	5.3514	0.8124	2.9429	3.1170
	Accuracy	99%	17%	42%	93%	7%	22%	72%	29%	30%
CNN2 (data augmentation)	Ave_loss	0.5015	0.4993	0.5035	0.8005	0.7920	0.8111,	1.8655	1.8632	1.8598
	Accuracy	85%	85%	84%	71%	71%	71%	33%	33%	34%
CNN2 (without data augmentation)	Ave_loss	0.1006	8.3428	5.8147	0.3594	10.0771	6.3615	1.5459	2.3862	2.4464
	Accuracy	97%	16%	33%	87%	2%	19%	45%	24%	23%
Rot-CNN	Ave_loss	0.8478	0.8478	2.6548	0.6145	0.6145	3.7220	1.7946	1.7946	2.0903
	Accuracy	72%	72%	43%	79%	79%	33%	36%	36%	26%

	MNIST	FashionMNIST	CIFAR10	
CNN1	1,044,096	1,044,096	1,437,888	
CNN2	25,440	25,440	98,656	
Rot-CNN	25,440	25,440	98,656	

	MNIST	FashionMNIST	CIFAR10
CNN1	193	184	206
CNN2	17	18	21
Rot-CNN	18	19	20

參數量

每個epoch所需運行的時間(s)

討論

- a. 在Rot-CNN中,對於Test與rot90x的結果始終保持一致。這是可以預期的,因為經過對稱filter的運算後,即使模型未看過rot90x的資料,但因為算出來的結果一致,會得到與Test資料集相同的數值。
- b. Rot-CNN在rot90x與ranrot的表現,除了在CNN1的ranrot表現小幅落後以外,對於沒做擴增資料集的CNN1、CNN2皆為領先。
- c. Rot-CNN所需訓練的參數量大幅少於CNN1,存在40倍的差距。 這也反映在每個epoch所需訓練的時間上,兩者相差近十倍。
- d. 儘管Rot-CNN對於旋轉有一定的抗性,但在Test的表現上顯著落後其他CNN的表現。

結論

新模型Rot-CNN對filter的限制的確對旋轉產生了抗性。 透過對filter的限制,附帶達成減少參數和提高速度的效果,但同時也要承受對本身資料集辨識度下降的代價。





題介紹影片