

Documents multimédia : description et recherche automatique

M2P GI, examen du 27 mai 2020, 2 heures, calculatrice et documents autorisés.

Les réponses doivent être justifiées de manière concise.

Exercice 1. Filtrés de convolution

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Image I

-1	1	-1
2	1	3
5	-2	4

Filtre F1

+1/2	0	-1/2
------	---	------

Filtre F2

+1
-2
+1

Filtre F3

On veut convoluer l'image I avec les filtres F1, F2 et F3.

Question 1.1 : Si l'on effectue une convolution de I sans "padding", quelle sera la taille de l'image résultat avec les filtres F1, F2 et F3 ?

Question 1.2 : Même question avec un "padding" de (1,1) pour F1, de (1,0) pour F2, et de (1,1) pour F3.

Question 1.3 : Dans le cas sans "padding", donner la valeur du pixel en haut à gauche pour l'image I convoluée, respectivement avec les filtres F1, F2 et F3.

Question 1.4 : Combien d'opérations sont-elles nécessaires pour calculer un élément de l'image de sortie dans la convolution avec le filtre F1 ?

Question 1.5 : À quel genre d'opérateur correspondent les filtres F2 et F3 ?

Exercice 2. Histogrammes

0	2	3	3	3	4
1	2	2	3	4	4
1	3	3	4	2	4
5	5	4	2	2	4
3	2	2	4	4	4
0	1	3	3	3	4

Image

Question 2.1 : Calculer l'histogramme en niveau de gris de l'image ci-dessus.

Question 2.2 : Quelle est la taille d'un descripteur histogramme de couleurs sur des images RGB avec 8 "bins" par composantes de couleur.

Question 2.3 : Même question avec 5 "bins" par composantes de couleur et un histogramme par blocs avec un découpage 3×3 de l'image.

Question 2.4 : Quelles distances sont les plus appropriées pour comparer des histogrammes ?

Exercice 3. Apprentissage profond (deep learning)

On considère une couche de convolution qui prend en entrée 128 plans de taille 56×56 et qui produit en sortie 256 plans de la même taille avec des filtres de taille 3×3 .

Question 3.1 : Faut-il prévoir un padding ? Si oui, de quelle taille ?

Question 3.2 : Combien y a-t-il de paramètres dans la fonction de convolution ?

Question 3.3 : Combien d'opérations flottantes sont effectuées pour chaque image d'entrée ?

On considère une couche complètement connectée à 4096 entrées et 1000 sorties

Question 3.4 : Combien y a-t-il de paramètres dans la fonction correspondante ?

Question 3.5 : Combien d'opérations flottantes sont effectuées pour chaque image d'entrée ?

Question 3.6 : En comptant seulement la couche de convolution et la couche linéaire ci-dessus, combien faut-il d'opérations flottantes pour traiter un lot ("batch") de 128 images ?

Question 3.7 : À quoi servent les couches "ReLU" dans les réseaux de neurones ? Par quoi peut-on les remplacer ?

Question 3.8 : À quoi servent les couches "Max Pool" dans les réseaux de neurones ? Par quoi peut-on les remplacer ?

Question 3.9 : À quoi sert la fonction de coût ou de perte ("Loss") dans les réseaux de neurones ? Donnez deux exemples de telles fonctions.