

INF8770

Technologies multimédias

A2022 - Travail pratique #3

Description du contenu pictural

Remise du travail :

- Au plus tard, le 5 décembre 2022, 8h30 sur Moodle - **aucun retard accepté**

Documents à remettre :

- Votre code source ainsi qu'un rapport en **PDF** dans une archive (.zip/.7z/...).
- Vous devez remettre votre rapport sur www.turnitin.com. Les instructions sont sur Moodle.

Autres directives :

- Il vous est interdit d'utiliser du code écrit par d'autres équipes.
- Il vous est permis d'utiliser du code trouvé sur internet et des bibliothèques externes, mais vous devez nous donner toutes les références.
- Le code lui-même n'est pas évalué, mais il doit être remis au complet.

Pénalités éventuelles :

- Trop de chiffres significatifs dans les tableaux/figures (**-0.5/20 point**);
 - Non remise du rapport sur Turnitin (**-0.5/20 point**);
 - Le rapport n'est pas sous le format pdf (**-0.5/20 point**);
 - Le document remis sur Moodle ne correspond pas exactement à celui remis sur Turnitin (**-0.5/20 point**);
 - Référence non citée (**-0.5/20 point**)
-

Préambule :

Ce travail pratique inclut une évaluation des qualités 2 et 5 du Bureau canadien d'agrément des programmes de génie (BCAPG).

- **Qualité 2 Analyse de problèmes** : capacité d'utiliser les connaissances et les principes appropriés pour identifier, formuler, analyser et résoudre des problèmes d'ingénierie complexes et en arriver à des conclusions étayées.
- **Qualité 5 Utilisation d'outils d'ingénierie** : capacité de créer et de sélectionner des techniques, des ressources et des outils d'ingénierie modernes et de les appliquer, de les adapter et de les étendre à un éventail d'activités simples ou complexes, tout en comprenant les contraintes connexes.

Travail à réaliser :

Dans ce travail vous devez implémenter un algorithme qui permet la description du contenu pictural : À partir d'une image de requête et d'une banque d'images dont le nom du fichier décrit l'objet présent dans l'image, vous devez retrouver quel objet est représenté dans l'image de requête.

Objectif de l'algorithme

Étant donnée une image de requête, vous devez trouver sa catégorie d'objets à partir de la banque d'images à laquelle l'image requête correspond le mieux. Pour ce faire, vous devez implémenter un algorithme qui mesure la distance, comme vu en cours, entre le contenu de l'image requête et des images de la banque d'images. La sortie de l'algorithme produit les 3 meilleures correspondances (top-3) sur la base de la distance sachant qu'on peut inclure une distance seuil au delà de laquelle on considère qu'une image n'a aucune correspondance. Sur la base de cette sortie, déduisez l'objet qui représente le mieux l'image requête. La figure 1 illustre l'approche générale de l'algorithme.

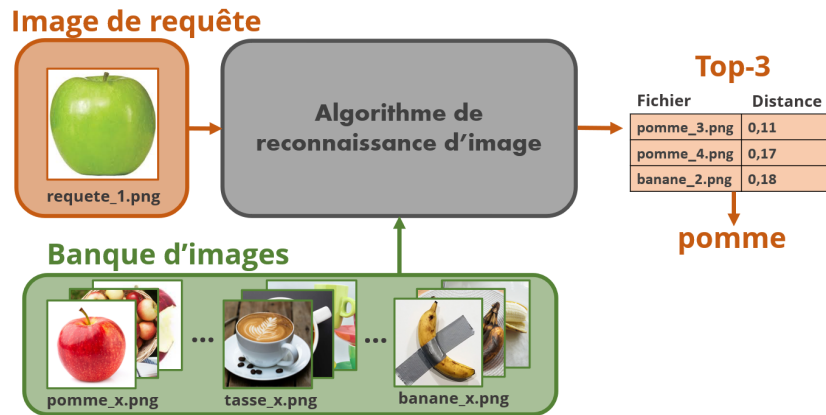


FIGURE 1 : Illustration de l'algorithme de reconnaissance d'image

Description des données

Les données sont disponibles sur Moodle et sont réparties comme suit : La banque d'images dans le dossier *banque d'images* contient 6 différentes catégories d'objets avec 5 différentes images par catégories. Les fichiers sont nommés avec le nom de la catégorie en premier puis un chiffre pour l'identifier : *categorie_id.png*. Les images de requête sont quant à elles contenues dans le dossier *images de requête* et il y en a 8. Toutes les images sont sous format *.png* mais leurs dimensions peuvent différer. Il faudra alors expliquer comment vous traiter cette différence de dimensions.

Les catégories réelles des images de requête sont les suivantes.

requete_1.png	pomme
requete_2.png	pomme
requete_3.png	tasse
requete_4.png	zebre
requete_5.png	banane
requete_6.png	planche de surf
requete_7.png	snowboard
requete_8.png	Inconnu

TABLE 1 : Catégories réelles associées aux images de requête

Évaluation Sachant que votre algorithme produit les 3 meilleures correspondances, on vous demande d'appliquer un vote majoritaire pour déterminer à quelle catégorie l'image se rapporte le plus selon votre algorithme.

Le vote majoritaire se calcule en prenant l'élément qui apparaît le plus souvent parmi les 3 meilleures correspondances. Si aucun n'apparaît plus souvent, sélectionnez celui qui a la plus petite distance.

Pour mesurer la performance de votre algorithme, vous utilisez la métrique de précision donnée ci-dessous qui mesure le taux de valeurs positives sur le nombre total de valeurs retrouvées.

$$Precision = \frac{|\text{Images correctement classées}|}{8} \quad (1)$$

Question 1 (/5)

Alice et Bob sont deux personnes étudiantes n'ayant pas suivi le cours INF8770. Ils proposent cependant deux algorithmes permettant de décrire le contenu d'une image.

Alice propose de calculer une distance euclidienne entre les pixels de l'image requête avec toutes les images de la banque de données. Dès que la distance est nulle, elle retourne l'image en question ainsi que la catégorie de l'objet auquel elle correspond. Si la distance nulle n'est jamais atteinte, elle renvoie les trois images avec la distance la plus petite, ainsi que la catégorie de l'objet identifié.

Bob propose de décrire les images de la banque de données par un histogramme de couleurs (un fichier d'indexation additionnel peut être utilisé). Puis, étant donnée une image requête, Bob calcule son histogramme de couleurs et retourne les trois images de la banque de données les plus proches en termes d'histogramme de couleurs avec la distance euclidienne, ainsi que la catégorie de l'objet identifié.

D'après vos connaissances en technologies multimédias, discutez de la pertinence de chaque algorithme lors de l'application sur les images requêtes.

- **Qualité évaluée : 2.1 Identifier et formuler un problème**

Critère d'évaluation : Analyse des algorithmes présentés par rapport aux deux mesures de performance.

Question 2 (/5)

Implémentez l'un des deux algorithmes proposés et appliquez-le sur les images du dossier test. Indiquez le temps moyen d'exécution par image ainsi que les performances sur les images requêtes. Est-ce que les résultats en termes de performance sont conformes à vos remarques pour la question 1 ? Selon vous, est-ce que se baser uniquement sur la meilleure correspondance aurait donné de meilleurs résultats ? Et si on avait pris les 5 objets les plus proches ?

- **Qualité évaluée : 5.3 Créer ou adapter un outil**

Critère d'évaluation : Capacité à créer un code informatique fonctionnel.

- **Qualité évaluée : 5.4 Intégrer des outils**

Critère d'évaluation : Capacité à bien intégrer les fonctions de bibliothèques de traitement d'images dans son propre code.

Question 3 (/10)

Implémentez votre propre algorithme de reconnaissance d'images, différent de ceux proposés par Alice et Bob. Expliquez votre méthodologie. Discutez de vos choix d'implémentation. Indiquez le temps moyen d'exécution par image ainsi que les performances sur les images requêtes. Comparez vos mesures de performances avec la méthode implémentée à la question 2. Synthétisez vos résultats dans un tableau ou un graphique. Discutez des résultats et des limitations de votre algorithme.

- **Qualité évaluée : 2.2 Explorer des approches de résolution et planifier la démarche**

Critère d'évaluation : Justification des choix de votre algorithme au regard des limites rencontrées par ceux décrits à la question 1.

- **Qualité évaluée : 2.4 Produire des résultats**
Critère d'évaluation : Pertinence des résultats et du protocole pour démontrer la qualité de la solution proposée.
- **Qualité évaluée : 2.5 Valider ses résultats et recommander**
Critère d'évaluation : Qualité et exhaustivité de l'analyse des résultats.
- **Qualité évaluée : 2.6 Analyser l'incertitude, la sensibilité et les limites des approches**
Critère d'évaluation : Qualité des explications sur les limites de la méthode.