

Traitement d'images

Réalisation d'un système de classification des panneaux routiers

Les panneaux routiers et les feux de signalisation et autres dispositifs de circulation sont utilisés pour avertir, guider ou informer les usagers de la route. Les panneaux routiers sont conçus pour être facilement reconnus par les conducteurs, principalement parce que leurs formes et leurs couleurs se distinguent facilement de leur environnement. Ces caractéristiques discriminantes permettent de classer les panneaux en plusieurs classes, selon leurs formes et leurs couleurs. Les panneaux sont classifiés en cinq catégories :

Les panneaux d'avertissement de danger : ils indiquent un danger sur la route. Ils se caractérisent par un triangle équilatéral avec un bord rouge épais et un intérieur blanc ou jaune. Un pictogramme dedans sert à spécifier différents avertissements (figure 1).



Figure 1 : Panneaux d'avertissement de danger

Les panneaux d'interdiction ou de restriction : ils signalent à l'utilisateur de la route qu'il doit respecter l'interdiction indiquée. Ils sont circulaires avec un fond blanc, jaune ou bleu, bordé d'une couronne rouge (figure 2).



Figure 2 : Panneaux d'interdiction ou de restriction

Les panneaux de priorité : ils sont destinés à notifier ou à porter à la connaissance des usagers de la route des règles particulières de priorité à des intersections (figure 3).



Figure 3 : Panneaux de priorité

Traitement d'images

Les panneaux d'obligation : sont des panneaux de positions, placés à l'endroit même où ils obligent à effectuer une action. Ils peuvent aussi indiquer des voies réservées à certains usagers. Ils se caractérisent par une forme circulaire, un fond bleu et un pictogramme blanc (figure 4).



Figure 4 : Panneaux d'obligation

Les panneaux de prescriptions particulières : ils sont généralement carrés ou circulaires, à fond bleu avec un symbole ou une inscription de couleur claire, ou à fond clair avec un symbole ou une inscription de couleur foncée (figure 5).



Figure 5 : Panneaux de prescriptions particulières

Ensemble de données utilisé

Dans ce projet l'ensemble de données GTSRB¹ (**German Traffic Sign Recognition Benchmark**) est adoptée. Cependant, uniquement 3 types de panneaux sont considérés, à savoir, les panneaux d'avertissement de danger, les panneaux d'interdiction ou de restriction et les panneaux d'obligation. La figure 6 montre quelques exemples extraits de cet ensemble de données. Chaque type de panneaux contient 50 images.



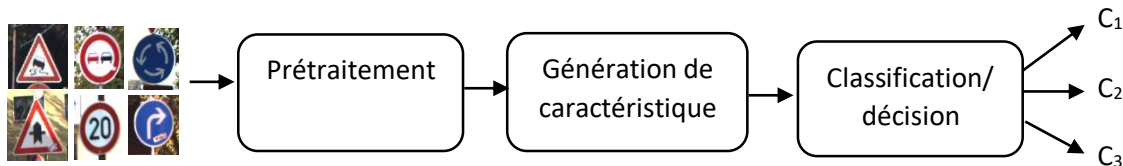
Figure 6 : Exemples de panneaux de l'ensemble de données GTSRB

¹<https://paperswithcode.com/dataset/gtsrb>

Traitement d'images

Description de l'approche

Un système de reconnaissance de forme est généralement composé de 3 modules essentiels après l'acquisition des données ; un module de prétraitement, un module de génération de caractéristique et un module de classification/décision.



1. **Prétraitement** : L'étape de prétraitement est une étape importante mais pas toujours indispensable. Elle vise à améliorer la qualité de l'image en améliorant son contraste.
2. **Génération de caractéristiques** : La caractérisation d'images est la transformation des données brutes en mesures plus informatives dites «attributs» générant ainsi ce qu'on appelle «vecteur caractéristique». Afin de caractériser au mieux les panneaux routiers, deux de leurs caractéristiques sont pris en compte : **la couleur et la forme**.

- a. **Caractérisation de la couleur (modélisation colorimétrique)** : La détection des panneaux de signalisation peut être réalisée en fonction de leur couleur. Les images sont converties du modèle Rouge, Vert et Bleu (RVB) vers le modèle Y'CbCr.

Ce modèle est une manière de représenter l'espace colorimétrique en vidéo, issue essentiellement des problèmes de transmission hertziennne. Une image captée par n'importe quel appareil est la somme des couleurs qui la composent. Le modèle Y'CbCr est formé de la composante Y' qui représente la luminance, plus deux informations de chrominance, Cb (Y' moins le bleu) et Cr (Y' moins le rouge).

Etant donné un pixel numérique représenté au format RGB, 8 bits, les composantes Y'CbCr peuvent être obtenues selon les équations suivantes :

$$Y' = 0.299R + 0.587G + 0.114B$$

$$C_b = -0.1687R - 0.3313G + 0.5B + 128$$

$$C_r = 0.5R - 0.4187G - 0.0813B + 128$$

- b. **Caractérisation de la forme : Histogramme du gradient orienté**

L'histogramme du gradient orienté (en anglais, histogram of oriented gradients ou HOG) est une caractéristique utilisée en vision par ordinateur pour la détection d'objet. L'idée importante derrière le descripteur HOG est que l'apparence et la forme locale d'un objet dans une image peuvent être décrites par la distribution de l'intensité du gradient ou la direction des contours. Ceci peut être fait en divisant l'image en régions adjacentes de petite taille, appelées cellules, et en calculant pour chaque cellule l'histogramme des directions du gradient ou des orientations des contours pour les pixels à l'intérieur de cette cellule. La combinaison des histogrammes forme alors le descripteur HOG.

3. **Classification** : Cette phase est le noyau de la reconnaissance des formes. Il s'agit d'associer une classe C à un vecteur caractéristique (forme inconnue), en utilisant les modèles (paramètres) obtenus lors

Traitement d'images

de l'apprentissage. Son principe : On dispose d'un ensemble de formes dont la classe est connue (base d'apprentissage ou de références) et on met au point une méthode qui, en étudiant cette base, sera ensuite capable de classer des formes inconnues. Dans ce travail 2 classifieurs sont abordés ; k-plus proches voisins et le réseau de neurones multicouche.

Travail demandé

Pour l'ensemble des images des 3 classes :

1. Appliquer un filtrage médian.
2. Réaliser une conversion de l'espace RGB vers l'espace Y'CbCr
3. Redimensionner les images Y'CbCr obtenues de façon à avoir des images de 64x128 pixels.
4. Générer l'histogramme des gradients orientés avec les paramètres 9 orientations, cellules de 8x8 pixels et blocs de 16x16 pixels. Quelle est la taille du vecteur caractéristique obtenu pour chaque image? Justifier.
5. Diviser les données de chaque classes aléatoirement en 2 ensembles ; 2/3 pour l'ensemble apprentissage (référence) et 1/3 pour l'ensemble test.
6. Réaliser une classification avec l'algorithme du k-plus proches voisins
 - a. Faire une étude sur l'influence de la distance choisie (Euclidienne, Manhattan et Canberra) en fonction de k (k=1,3,5,7 et 9)
 - b. Analyser la matrice de confusion pour le meilleur résultat obtenu.
7. Réaliser une classification à l'aide d'un réseau de neurones multicouche (faire le tuning des paramètres de façon à obtenir une précision de 95% sur les données apprentissage avant de passer aux données test).
 - a. Analyser la matrice de confusion pour la meilleure précision obtenue.
8. Comparer les deux méthodes de classification ; K-ppv et RNM
9. Quelles sont les images qui ont été mal classées ? Justifier
10. Faire des propositions/ suggestions pour améliorer les performances atteintes.

Distance	Equation
Euclidienne	$D_{euclidienne}(a, b) = \sqrt{\sum_{i=1}^n a_i - b_i ^2}$
Manhattan	$D_{Manhattan}(a, b) = \sum_{i=1}^n a_i - b_i $
Canberra	$D_{Canberra}(a, b) = \sum_{i=1}^n \frac{ a_i - b_i }{ a_i + b_i }$

RNM sur Python

Traitement d'images

https://scikit-learn.org/stable/modules/neural_networks_supervised.html

Histogramme des gradients orientés

Histogram of Oriented Gradients - HOG

Le choix d'un descripteur est crucial pour la performance de n'importe quel système de reconnaissance de formes : il doit être le plus représentatif, le plus discriminant et le plus rapide en termes de temps de calcul. Ainsi, un bon descripteur doit être capable de capturer les similitudes essentielles entre les objets de la même classe et les différences avec des objets des classes concurrentes.

L'histogramme des gradients orientés (HOG) est un descripteur utilisé dans de nombreuses applications de vision par ordinateur et de traitement d'images à des fins de détection d'objets. Il a été proposé pour la détection de personne en 2005, par Dalal et Triggs chercheurs à l'INRIA de Grenoble.

L'idée importante derrière le descripteur HOG est que l'apparence et la forme locale d'un objet dans une image peuvent être décrites par la distribution de l'intensité du gradient ou la direction des contours. Ceci peut être fait en divisant l'image en régions adjacentes de petite taille, appelées **cellules**, et en calculant pour chaque cellule l'histogramme des directions du gradient ou des orientations des contours pour les pixels à l'intérieur de cette cellule. La combinaison des histogrammes forme alors le descripteur HOG. Pour de meilleurs résultats, les histogrammes locaux sont normalisés en contraste, en calculant une mesure de l'intensité sur des zones plus larges que les cellules, appelées des **blocs**, et en utilisant cette valeur pour normaliser toutes les cellules du bloc. Cette normalisation permet une meilleure résistance aux changements d'illuminations et aux ombres.

La génération de l'histogramme du gradient orienté pour une image passe par plusieurs étapes :

1- Calcul du gradient

La première étape est le calcul du gradient en appliquant un filtre dérivatif 1-D centré, dans les directions horizontales et verticales en utilisant les masques suivants:

$$[-1,0,1] \text{ et } [-1,0,1]^T$$

Par la suite, l'amplitude ainsi que l'orientation du gradient sont calculés à travers les équations suivantes :

$$|G| = \sqrt{G_x^2 + G_y^2} , \theta = \tan^{-1} \left(\frac{G_y}{G_x} \right)$$

Dans le cas des images couleurs, le gradient est calculé séparément pour chaque composante, et on retient pour chaque pixel le gradient de plus grande norme.

Un exemple de calcul du gradient sur une cellule de 8x8 pixels est montré dans la figure 1.

Traitement d'images

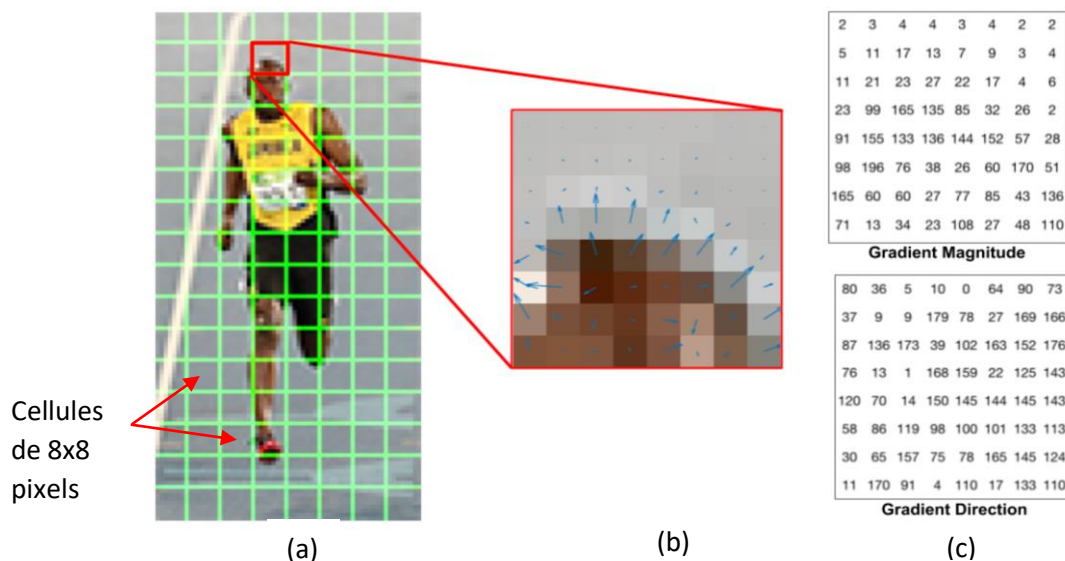


Figure 1 : Exemple de calcul du gradient

La figure 1 –(a) représente une image couleur qui est divisée en cellule de 8x8 pixels. La figure 1-(b) montre la cellule de l'image superposée de flèches représentant le gradient - la flèche montre la direction du gradient et sa longueur montre son amplitude. A noter que la direction des flèches indique la direction du changement d'intensité et l'amplitude montre l'ampleur de la différence. La figure 1-(c) montre l'amplitude et la direction du gradient dans la cellule de 8x8 pixels. Les directions sont comprises entre 0 et 180 degrés au lieu de 0 à 360 degrés. Ces gradients sont appelés « non signés » car un gradient et son négatif sont représentés par les mêmes nombres. En d'autres termes, une flèche de gradient et son opposée à 180 degrés sont considérées comme identiques. Mais pourquoi ne pas utiliser le 0 – 360 degrés ? Empiriquement, il a été démontré que les gradients non signés fonctionnent mieux que les gradients signés pour la détection des piétons. Certaines implémentations de HOG vous permettront de spécifier si vous souhaitez utiliser des gradients signés.

2- Formation de l'histogramme

Dans cette étape, l'image est divisée en **cellules** de 8x8 pixels et un histogramme de gradients est calculé pour chaque cellule (Le choix de 8x8 pixels est un choix de design qui peut être adapté selon le besoin). L'histogramme est essentiellement un vecteur de 9 cases (nombres) correspondant aux angles 0°, 20°, 40°, 60° ... 160°. Reprenons la cellule de 8 × 8 pixels précédente (figure 1). La figure 2 illustre le processus de formation de l'histogramme.

Une case est sélectionnée en fonction de la direction du gradient et le vote (la valeur qui entre dans la case) est sélectionné en fonction de l'amplitude du gradient. Sur la figure 2, le pixel entouré de bleu a un angle (direction) de 80° et une amplitude de 2. On ajoute donc 2 à la 5^{ème} case de l'histogramme. Le gradient au niveau du pixel entouré de rouge a un angle de 10° et une amplitude de 4. Étant donné que 10° se situe à mi-chemin entre 0 et 20, le vote du pixel se répartit uniformément dans les deux cases.

Traitement d'images

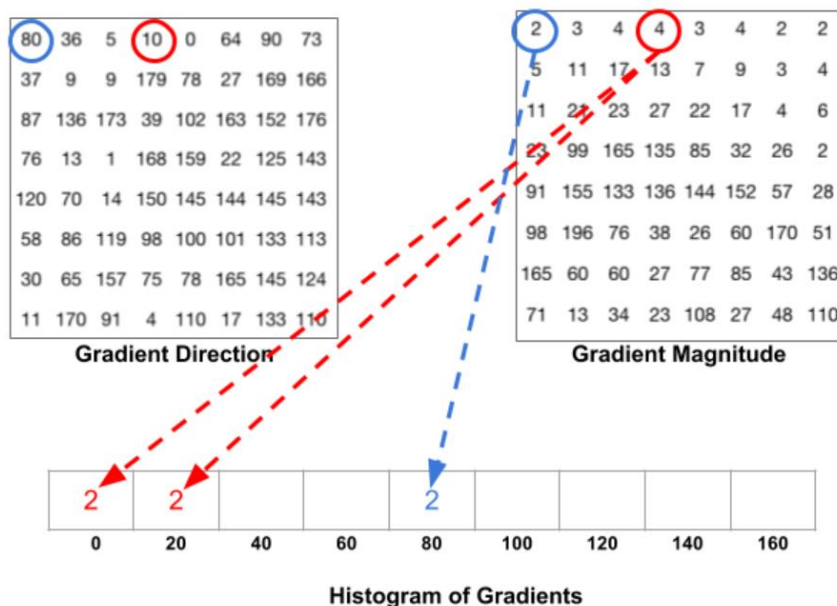


Figure 2 : Formation de l'histogramme

Si l'angle est supérieur à 160° et est compris entre 160° et 180° , le dit pixel contribue proportionnellement à la case 0° et 160° . Ainsi, dans l'exemple ci-dessous, le pixel avec un angle de 165° contribue proportionnellement aux cases 0° et 160° .

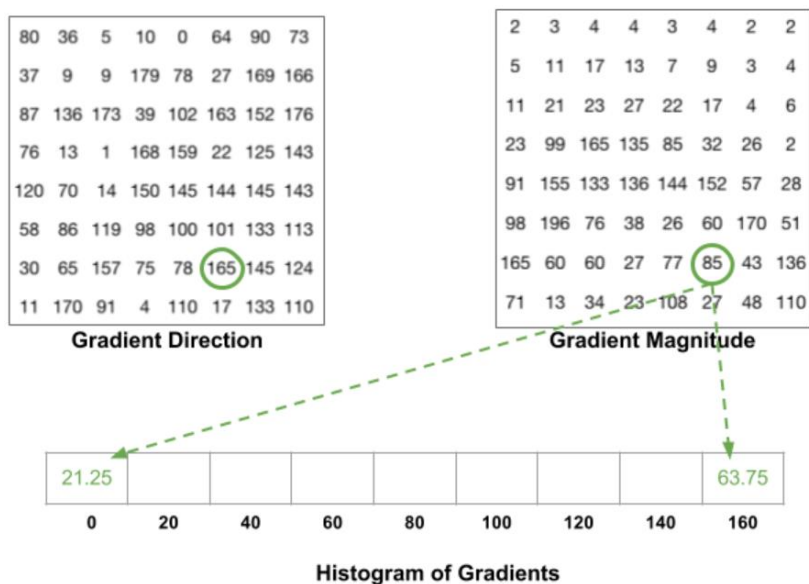


Figure 3 : Cas où l'orientation du gradient est entre 0° et 180°

Les contributions de tous les pixels des cellules 8×8 sont additionnées pour créer l'histogramme à 9 cases. Pour la cellule précédente (figure 1), l'histogramme aura l'allure suivante :

Traitement d'images

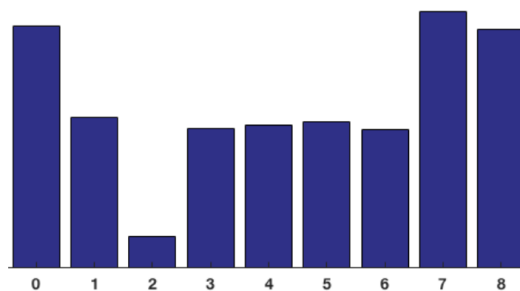


Figure 4 : Histogramme des gradients orientés à 9 directions

3- Normalisation par bloc

Dans l'étape précédente, nous avons créé un histogramme basé sur le gradient de l'image. Le gradient d'une image est sensible à l'éclairage global. Si l'image est assombrie en divisant toutes les valeurs de pixels par 2, l'amplitude du gradient changera de moitié et, par conséquent, les valeurs de l'histogramme changeront de moitié. Idéalement, un descripteur doit être indépendant des variations d'éclairage. En d'autres termes, l'histogramme doit être « normalisé » afin qu'il ne soit pas affecté par les variations d'éclairage.

Avant d'expliquer comment l'histogramme est normalisé, voyons comment un vecteur de longueur 3 est normalisé : soit un vecteur de couleur RGB [128, 64, 32]. La norme L2 de ce vecteur est $\sqrt{128^2 + 64^2 + 32^2} = 146,64$. La division de chaque élément de ce vecteur par 146,64 nous donne un vecteur normalisé [0,87, 0,43, 0,22]. Considérons maintenant un autre vecteur dans lequel les éléments sont deux fois la valeur du premier vecteur $2 \times [128, 64, 32] = [256, 128, 64]$. La normalisation de ce vecteur donnera [0.87, 0.43, 0.22], ce qui est identique à la version normalisée du vecteur RGB d'origine. On peut voir que la normalisation d'un vecteur supprime l'échelle.

Maintenant que nous savons comment normaliser un vecteur, nous pouvons être tentés de penser que lors du calcul de HOG, nous pouvons simplement normaliser l'histogramme 9×1 de la même manière que nous avons normalisé le vecteur 3×1 ci-dessus. Ce n'est pas une mauvaise idée, mais une meilleure idée est de normaliser sur un **bloc** de plus grande taille de 16×16 pixels.

Un bloc 16×16 pixels a 4 histogrammes qui peuvent être concaténés pour former un vecteur d'élément 36×1 et il peut être normalisé de la même manière qu'un vecteur 3×1 est normalisé. La fenêtre est alors déplacée de 8 pixels (voir figure 5) et un vecteur normalisé 36×1 est calculé sur cette fenêtre et le processus est répété.

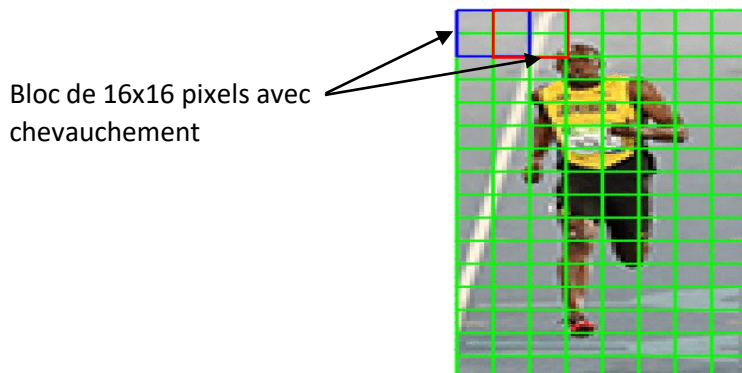


Figure 5 : Bloc de 16x16 pixels

Traitement d'images

4- Calcul du vecteur caractéristique HOG

Pour calculer le vecteur de caractéristique final pour l'ensemble des cellules de l'image, les vecteurs 36×1 de chaque bloc de 16×16 pixels sont concaténés en un vecteur géant.

Fonction HOG sur Python :

```
from skimage.feature import hog

fv, hog_image = hog(resized_img, orientations=9, pixels_per_cell=(8, 8),
                    cells_per_block=(1, 1), visualize=True, multichannel=True)
```

La fonction `hog()` prend 6 paramètres en entrée:

- **image**: l'image cible sur laquelle HOG va être calculé
- **orientations**: Nombre de directions dans l'histogramme à considérer (choisir 9 orientations)
- **pixels_per_cell**: Détermine la taille de la cellule, 8×8 pixels.
- **cells_per_block**: Nombre de cellules par bloc, 2×2 pour obtenir 16×16 pixels.
- **visualize**: Booléen ; si retourner l'image de HOG, mettre **True** pour pouvoir visualiser l'image gradient.
- **multichannel**: Mettre **True** pour dire à la fonction que la dernière dimension est considérée comme canal de couleur.

La fonction retourne en sortie 2 variables :

- **fv** : vecteur caractéristique de l'histogramme des gradient orientés
- **hog_image** : image gradient