

TD 2 – MÉTRIQUES D'ATTENTION VISUELLE

La notion de saillance visuelle fait référence à la propriété visuelle faisant que certains objets de notre entourage soient plus perçus que d'autres, ils « capturent notre attention ».

Il est très important pour un système biologique complexe (humains, animaux) d'être capable de détecter rapidement de potentielles proies, prédateurs, ou autres dans un monde rempli d'information visuelle. Primates ou autres animaux filtrent cette information visuelle afin de concentrer leur processus inconscient de reconnaissance d'objets sur un nombre restreint d'objets dits « saillants ».

De nombreux travaux en vision par ordinateur s'inspirent de cette propriété visuelle afin de concentrer leurs calculs sur les régions saillantes. Dans la figure suivante, il est aisé de voir que pour un problème de reconnaissance d'action (ici « faire un café »), les objets significatifs à reconnaître sont la cuillère et le pot de café en poudre.



La figure de gauche montre la frame d'origine. La figure de droite montre une « carte de chaleur » de valeurs de saillance, appelée « carte de saillance », calculée automatiquement à partir d'un algorithme exploitant les propriétés de l'œil.

Sur cet exemple, il est visuellement aisé de conclure que l'algorithme automatique d'extraction de saillance montre de très hautes performances. Cependant ce n'est peut-être pas forcément le cas pour toutes les images de la vidéo. Et comment évaluer si cet algorithme donne des performances au dessus des autres ?

De nos jours, la prédiction de la saillance dans les images et la vidéo par les méthodes d'apprentissage de la machine est un sujet de recherche intensif. Les réseaux de neurones convolutifs profonds (CNNs profonds) se sont révélés être un puissant moyen de prédiction des zones saillantes. En ajoutant des informations de mouvement résiduel à l'entrée de CNNs profonds, l'architecture du réseau ChaboNet développé dans le LABRI pour prédire la saillance en vidéo naturelle est résumée dans la figure suivante.

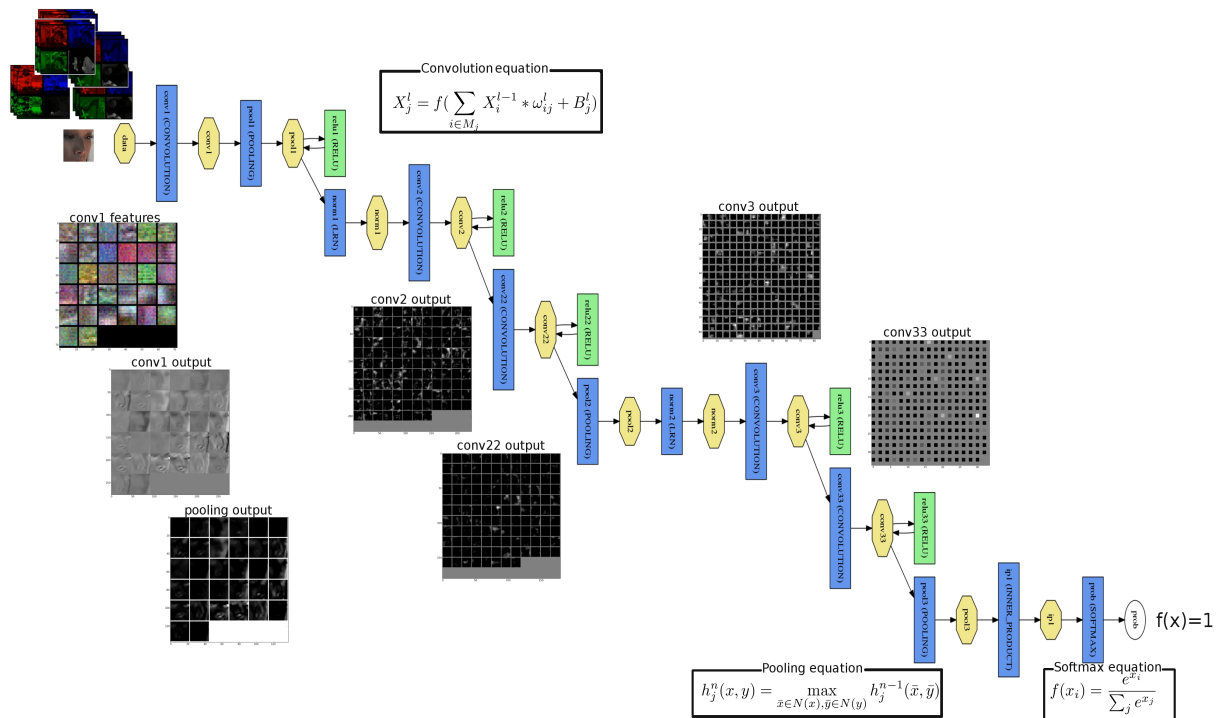


Figure: "ChaboNet" réseau CNN profond pour prédiction de la saillance dans la vidéo naturelle.

Le modèle du réseau «ChaboNet» est déjà défini et entraîné. L'étape d'entraînement nécessite un important temps de calcul et une capacité importante de la GPU. Pour simplifier, dans ce Td vous allez juste utiliser ce modèle «**ChaboNet**» pré-entraîné pour commencer la prédiction de la saillance dans certaines vidéos.

Le premier point d'intérêt: comment utiliser un modèle formé pour créer une carte de saillance dense.

Il est visuellement facile de conclure que la prédiction automatique de la saillance montre des performances très élevées. Cependant, ce n'est peut-être pas nécessairement le cas pour tous les images de la vidéo. Et comment évaluer si cet algorithme fournit des performances supérieures aux autres?

Le deuxième point d'intérêt : comment mesurer la performance d'un algorithme pour la création automatique de cartes de saillance. Vous allez calculer l'extraction algorithmique automatique de saillance visuelle à l'aide de trois métriques différentes vues en cours.

1- PRÉDICTION DE LA CARTE DE SAILLANCE AVEC DEEP CNN

Dans votre dossier de travail, créer un dossier « saliencyChabonet » qui va contenir la saillance prédites par le modèle proposé.

Ouvrir un terminal et se logger à /usr/local/bin/launcher/caffe_chroot

accéder au répertoire caffe_master situé dans /opt/caffe/

Exécuter le code python créé pour prédire la carte de saillance (ajouter les arguments nécessaires).

```
>> python ./python/ChaboNet4k_HOLLYWOOD_SaliencyGenerate.py
```

La prédiction de la saillance nécessite du temps pour être générée, pour ne pas perdre de temps laissez le programme tourner et passer à la deuxième section du TD.

2- RÉCUPÉRATION DE LA BASE DE DONNÉE ET DU CODE

Télécharger le dataset et code à l'adresse suivante :

http://dept-info.labri.fr/~schaabou/TD_saliency/section2/tm_td4_saillance_sec2.zip

Extraire l'archive dans votre dossier de travail.

Description de l'archive :

```
├── Dataset
│   ├── Automatic_ChaboNet_saliency
│   │   ├── actioncliptest_350
│   │   ├── actioncliptest_451
│   │   └── Original_frames_half
│   │       ├── actioncliptest_350
│   │       └── actioncliptest_451
│   └── Visual_Points
│       ├── actioncliptest_350
│       └── actioncliptest_451
├── gbvs
├── Matlab_student
│   ├── auc.m
│   ├── calcAUCscore.m
│   ├── calcCCscore.m
│   ├── calcNSSscore.m
│   ├── index_cliptest_350.txt
│   ├── index_cliptest_451.txt
│   ├── main.m
│   └── preview_saillance_et_fixations.m
```

Le dossier « dataset » est composé de trois dossiers contenant les trames des deux vidéos "actioncliptest_350" et "actioncliptest_451", les fixations d'œil enregistrées

(« Visual_Points »), et les cartes de saillances calculées automatiquement («Automatic_ChaboNet_saliency»).

Le dossier contenant le code et fichiers d'index se nomme « Matlab_étudiants ».

Le dossier « gbvs » servira à calculer de nouvelles cartes de saillance automatiques avec le modèle de Itti.

3- PREMIERS TESTS

Durant ce projet vous allez progressivement compléter le fichier principal « main.m ». Vous pouvez dès à présent l'ouvrir et le compiler (ce qui n'a pas d'effets visibles pour le moment).

Les fichiers servant à calculer les métriques sont déjà fournis («auc.m», « calcAUCscore.m», « calcCCscore.m», « calcNSSscore.m»).

Le fichier «preview_saillance_et_fixations.m », est fourni afin de vous fournir un aperçu graphique des cartes de saillances et points d'intérêts.

4- LIRE LE FICHIER D'INDEX

Le fichier "index_cliptest_350.txt" et "index_cliptest_451.txt" contient dans chaque ligne le chemin vers la saillance automatique, les points visuels, et l'image correspondante. C'est le lien avec la base de données qu'il vous faut dans un premier temps ouvrir et lire avec Matlab.

Une bonne méthode est d'ouvrir le fichier, stocker les données dans une variable (sous forme de cellules de « strings »), et fermer le fichier.

5- PREMIERS RÉSULTATS SUR UNE IMAGE

Dans un premier temps, avant de lancer les calculs de métriques sur toutes les cartes de saillances, calculer les scores AUC, PCC et NSS entre **la première image** de saillance et les fixations d'œil correspondantes.

Afin de se rendre compte de la valeur de ce score, calculez les mêmes scores entre ces fixations d'œil et une carte de saillance de mêmes dimensions contenant des valeurs **aléatoires** entières (entre 0 et 255)

Les valeurs théoriques obtenues pour ces métriques si un modèle est aléatoire sont :

- Pour AUC : 0.5
- Pour PCC : 0
- Pour NSS : 0

Comparez les scores obtenus entre ces valeurs et celles que vous obtenez pour le modèle aléatoire ainsi que la carte de saillance générée par algorithme.

6- BOUCLE PRINCIPALE ET RÉSULTATS

Créer maintenant la boucle principale servant à calculer les scores NSS, AUC et PCC entre toutes les cartes automatiques et fixations d'œil correspondantes dans la base de données.

Calculer et afficher les valeurs de moyenne et l'écart-type associés à chacune de ces métriques.

Utiliser la fonction *bar* pour afficher les résultats sous forme de 3 barres distinctes et la fonction *errorbar* pour afficher les barres d'erreur (l'écart-type) sur ces barres. Ne pas oublier d'utiliser *hold on* et *hold off*.

7- ET AVEC DES VALEURS ALÉATOIRES ?

Répéter l'étape 5 en remplaçant cette fois-ci chaque carte de saillance automatique par une carte à chaque fois générée aléatoirement (comme dans l'étape 4).

Calculer et afficher les valeurs de moyenne et l'écart-type associés à chacune des métriques.

Comparer avec les valeurs obtenues dans la partie 5 et avec les valeurs théoriques.

8- COMPARAISON AVEC L'ÉTAT DE L'ART

Une des méthodes automatiques de création de cartes de saillances la plus connue dans le domaine est celle de Laurent Itti. Le dossier «gbvs » contient la fonction *ittikichmap* permettant de calculer la saillance d'une image par cette méthode.

Ouvrez maintenant le fichier « calcul_itti_maps.m » et créez la boucle principale permettant de calculer et **enregistrer en .jpg** la saillance de Itti pour toutes les images de la vidéo.

Ré-ouvrir le fichier « main.m » et le compléter afin de calculer les scores et figures pour ces nouvelles cartes de saillance. Comparer es scores avec ceux calculés dans les parties 5 et 6.

9- RAPPORT

Il vous est demandé d'écrire un rapport concis contenant les parties suivantes:

-Introduction

-Présentation des résultats et interprétation (comment faire évoluer la performance et expliquer la raison).

-Conclusion (ou problèmes rencontrés ou résolus, perspectives possibles, ...)

Envoyez votre rapport, nommé MYNAME_TD2.tar.bz2, à souad.chaabouni@u-bordeaux.fr dans une archive qui contient également le code Matlab avant minuit du samedi 19 novembre 2016.