



Stage M2

**Compte Rendu Semaine 13**

Benjamin Serva  
Master 2 IMAGINE  
Université de Montpellier

18 avril 2025

Encadrants :

Olivier Strauss & William Puech & Frédéric Comby

# 1 Tâches effectuées cette semaine

## 1.1 Test des méthodes de segmentation sur les BDD

### 1.1.1 BDD 1 : BBC

Algo	Vrai Positif	Faux Positif	Précision	Rappel	F1 Score
Adaptive	3884	487	88.86%	80.00%	84.20%
Content	3609	1013	78.08%	74.34%	76.16%
Hash	3998	19 19	67.57%	82.35%	74.23%
Histogram	3710	2332	61.40%	76.42%	68.09%
Threshold	0	428	0.00%	0.00%	0.00%

Table 1: Comparaison des résultats des algorithmes pour la BDD BBC

Nombre total de plans à détecter : 4855.

### 1.1.2 BDD 2 : AutoShot

Algo	Vrai Positif	Faux Positif	Précision	Rappel	F1 Score
Adaptive	1116	951	53.99%	49.21%	51.49%
Content	1019	871	53.92%	44.93%	49.01%
Hash	856	827	50.86%	37.74%	43.33%
Histogram	905	1789	33.59%	39.90%	36.48%
Threshold	3	41	6.82%	0.13%	0.26%

Table 2: Comparaison des résultats des algorithmes sur la BDD AutoShot

Nombre total de plans à détecter : 2268.

## 1.2 Recherche de bases de code

Aucun des articles consultés ne propose de lien vers une implémentation des méthodes présentées. J'ai également recherché les profils des auteurs sur GitHub, mais aucun projet pertinent n'a été trouvé. Il faudra donc implémenter moi-même toutes les méthodes sélectionnées.

## 1.3 Réunion avec Monsieur Comby et Monsieur Puech

Nous avons discuté des résultats précédents ainsi que des prochaines étapes. Il a été décidé de se concentrer sur la détermination d'une image clé par plan, ainsi que sur les méthodes d'extraction d'image clé. Trois méthodes ont été sélectionnées que je dois désormais implémenter et tester.

## 1.4 Test des méthodes de segmentation sur les BDD de vidéo en noir et blanc

En calculant la segmentation sur un film en noir et blanc, j'ai constaté que la méthode considérée comme la plus performante précédemment ne fonctionnait pas du tout. J'ai donc converti les deux BDD en niveaux de gris pour relancer les tests. Les résultats sont les suivants :

### 1.4.1 BDD 3 : BBC Gray

Algo	Vrai Positif	Faux Positif	Précision	Rappel	F1 Score
Adaptive	1304	738	63.86%	26.86%	37.81%
Content	57	165	25.68%	1.17%	2.25%
Hash	3996	1915	67.60%	82.31%	74.23%
Histogram	3682	2408	60.46%	75.84%	67.28%
Threshold	1	412	0.24%	0.02%	0.04%

Table 3: Comparaison des résultats des algorithmes pour la BDD BBC (niveau de gris)

Nombre total de plans à détecter : 4855.

### 1.4.2 BDD 4 : AutoShot Gray

Algo	Vrai Positif	Faux Positif	Précision	Rappel	F1 Score
Adaptive	836	485	63.29%	36.86%	46.59%
Content	247	348	41.51%	10.89%	17.25%
Hash	860	821	51.16%	37.92%	43.56%
Histogram	939	1700	35.58%	41.40%	38.27%
Threshold	2	44	4.35%	0.09%	0.17%

Table 4: Comparaison des résultats des algorithmes pour la BDD Autoshot (niveau de gris)

Nombre total de plans à détecter : 2268.

## 1.5 Modification code du logiciel

Suite aux résultats précédents, j'ai modifié le code afin qu'il détecte automatiquement si le film est en noir et blanc ou en couleur, et qu'il adapte en conséquence la méthode de segmentation à appliquer.

## 1.6 Statistiques

Ces statistiques ont été réalisées sur cinq films fournis par Monsieur Le Bihan (La vie d'Adèle, MOF, Shining, Falstaff, 4°ECGTB) :

- Durée moyenne d'un film : 01:55:53.7
- Nombre moyen de frames par film : 172,370.00
- Nombre moyen de plans (détectés par la segmentation) : 1033.80
- Durée moyenne d'un plan : 00:00:06.7
- Nombre moyen de frames par plan : 166.73
- Médiane : 00:00:03.7
- 1er quartile : 00:00:02.0
- 3e quartile : 00:00:07.2
- Écart-type : 00:00:10.7

## 1.7 Première passe d'analyse

L'objectif de cette première passe est d'avoir une vue d'ensemble sur le contenu de la vidéo, afin de potentiellement adapter la méthode utilisée. Avec YOLOv8, il est possible d'identifier facilement des objets. Voici un exemple de résultats obtenus sur une vidéo d'une discussion dans un bar :

- person: 1028 frames  $\rightarrow$  98.09%
- wine glass: 207 frames  $\rightarrow$  19.75%
- cup: 163 frames  $\rightarrow$  0.1555%
- dining table: 161 frames  $\rightarrow$  15.36%
- bottle: 32 frames  $\rightarrow$  3.05%
- chair: 1 frames  $\rightarrow$  0.10%

## 2 Objectifs pour la semaine prochaine

- Implémenter la première méthode d'extraction d'une image clé par plan.
- La comparer à des méthodes naïves (ex. : prendre la 10ème frame ou la frame centrale du plan).