

Détection d'enluminures dans des manuscrits médiévaux IIF à l'aide de l'apprentissage profond

Victoria Eyharabide¹ Fouad Aouinti¹ Xavier Fresquet²

¹Laboratoire STIH, Sorbonne Université, France

²IReMus Laboratory, Sorbonne Université, France

24 mars 2021

Article soumis à ICDAR 2021 (16th International Conference on Document Analysis and Recognition)

Plan

- 1 Introduction
- 2 Notre proposition
- 3 Évaluation et discussion
- 4 Conclusion et perspectives

Plan

- 1 Introduction
- 2 Notre proposition
- 3 Évaluation et discussion
- 4 Conclusion et perspectives

Introduction



FIGURE 1 – Enluminure d'un caractère "B"

<https://iif.bodleian.ox.ac.uk/iif/canvas/9da1b4c2-2c04-40e9-a171-dc4b656af141.json>

Plan

- 1 Introduction
- 2 Notre proposition**
- 3 Évaluation et discussion
- 4 Conclusion et perspectives

Notre proposition

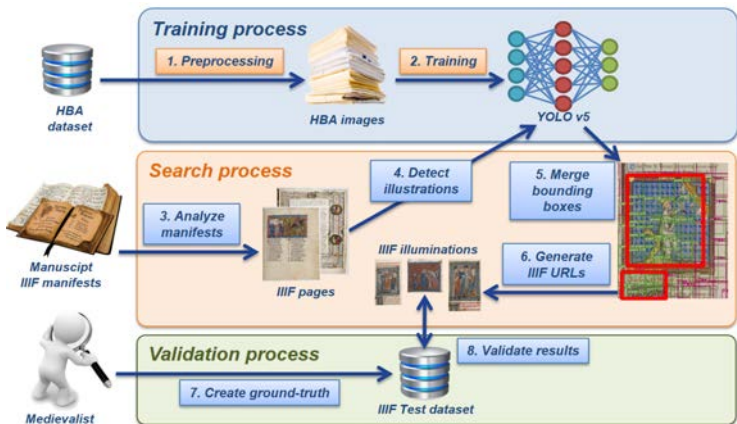


FIGURE 2 – Aperçu de notre proposition pour détecter des enluminures

Example



FIGURE 3 –

https://www.e-codices.unifr.ch/loris/ubb/ubb-F-II-0020/ubb-F-II-0020_054r.jp2/full/full/0/default.jpg

https://www.e-codices.unifr.ch/loris/ubb/ubb-F-II-0020/ubb-F-II-0020_054r.jp2/499,1790,702,661/full/0/default.jpg

Préparation du jeu de données HBA

- Le jeu de données HBA contient :
 - 4436 images vérité terrain ;
 - 1429 images annotées au niveau du pixel (six classes).



- Certaines spécificités du jeu de données HBA pourraient influencer considérablement le processus d'apprentissage :
 - dimensions larges et dispersées des images ;
 - hétérogénéité, style d'illustration, mises en page complexes, etc.

- Après l'étape de pré-traitement, nous avons obtenu 366 images annotées au niveau de pixel en utilisant trois classes (*graphic*, *text*, *the_rest*).

Entraînement de YOLO sur le jeu de données HBA

- Pour la formation de YOLO, les 366 images HBA sélectionnées doivent être étiquetées et annotées avec des cadres de délimitation entourant les enluminures.
 - Pour chaque image, nous avons généré aléatoirement 10 000 cadres de délimitation et leur avons attribué, en fonction du seuil de couleur d'arrière-plan, la classe appropriée.
 - Le sous-ensemble de formation contient 293 images vérité terrain (80% des images) et leurs fichiers d'annotations.
-
- Le jeu de données HBA préparé contient 94 084 annotations :
 - 31 181 → *graphic*;
 - 31 138 → *text*;
 - 31 765 → *the_rest*.

Fusion de boîtes englobantes

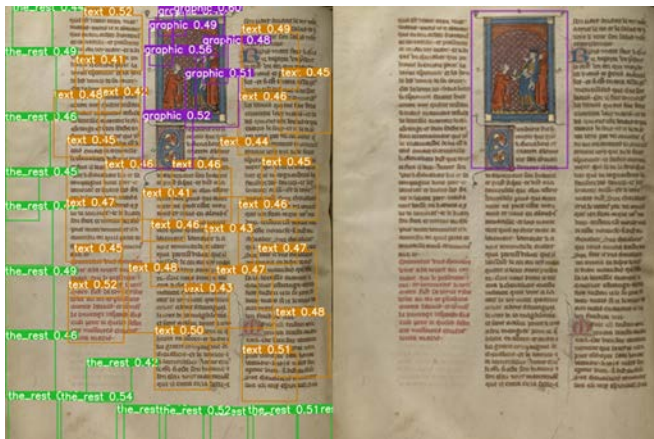


FIGURE 4 – Détection d'enluminure après la fusion des boîtes *graphic*

Génération d'URL IIIF à partir de boîtes englobantes

- IIIF Image API 3.0 spécifie une syntaxe d'URL normalisée pour la transmission d'une image via une requête HTTP de :
 - l'image elle-même (fichier image) ;
 - informations techniques sur l'image (fichier JSON).
- Le modèle d'URL pour demander une image doit être conforme à : `{scheme} ://{server}/{prefix}/{identifiant}/{region}/{size}/{rotation}/{quality}.{format}`
- La région est généralement spécifiée par la valeur *full*, ce qui signifie que l'image complète doit être renvoyée.
- Nous remplaçons la valeur par défaut de la région par les coordonnées de l'enluminure détectée par notre système en suivant la syntaxe spécifiée par le standard IIIF.

Plan

- 1 Introduction
- 2 Notre proposition
- 3 Évaluation et discussion**
- 4 Conclusion et perspectives

Création d'un jeu de données de test IIF

TABLE 1 – Caractéristiques du jeu de données de test IIF

Manuscrit	Valenciennes BM, 0099 (092)	Garrett MS. 28.	MS Nn.2.36	F II 20
Titre	Apocalypse	De animalibus	Gospel book	Bible
Date	800-900	1200- 1300	1120-1130	1270
Location	Valenciennes BnF	Universität Basel	University of Cambridge	Princeton University
Nombre de pages	100	316	489	930
Pages annotées	40	19	10	72

Résultats expérimentaux (1)

TABLE 2 – Résultats expérimentaux de la détection des enluminures

Bibliothèque	Images IIF	Images annotées	TP	TN	FP	FN	Précision	Rappel
BnF	100	40	37	46	14	3	0.73	0.93
Basel	316	19	16	287	10	3	0.62	0.84
Cambridge	489	10	9	461	18	1	0.33	0.9
Princeton	930	72	70	824	34	2	0.67	0.97



- La page originale de Basel :
https://www.e-codices.unifr.ch/loris/ubb/ubb-F-II-0020/ubb-F-II-0020_0031r.jp2/full/full/0/default.jpg
- L'enluminure détectée :
https://www.e-codices.unifr.ch/loris/ubb/ubb-F-II-0020/ubb-F-II-0020_0031r.jp2/664,2657,430,774/full/0/default.jpg

Résultats expérimentaux (2)



(a) Page avec règle, palette de couleurs (Basel)



(b) Couverture, petite décoration (Princeton)

FIGURE 5 – Exemples de résultats faux positifs

Plan

- 1 Introduction
- 2 Notre proposition
- 3 Évaluation et discussion
- 4 Conclusion et perspectives

Conclusion et perspectives

- Dans ce travail, nous avons formé YOLOv5 sur le jeu de données HBA pour détecter les enluminures dans les manuscrits médiévaux IIF.
- Les résultats obtenus sont satisfaisants et confirment les bonnes performances de l'approche proposée.
- Nous proposons d'utiliser les données de Biblissima pour tester le modèle sur un jeu de données beaucoup plus large.
- Nous envisageons aussi de mettre le standard IIF en service de la recherche en musicologie à l'aide des techniques d'apprentissage automatique.

Références



M. Mehri, P. Héroux, R. Mullot, J. P. Moreux, B. Coüasnon, B. Barrett.
HBA 1.0 : A pixel-based annotated dataset for historical book analysis.
4th International Workshop on Historical Document Imaging and Processing,
pp. 107-112, 2017.



S. A. Oliveira, B. Seguin, F. Kaplan.
dhSegment : A generic deep-learning approach for document segmentation.
16th International Conference on Frontiers in Handwriting Recognition
(ICFHR), pp. 7-12, 2018.



T. Monnier, M. Aubry.
docExtractor : An off-the-shelf historical document element extraction.
17th International Conference on Frontiers in Handwriting Recognition
(ICFHR), pp. 91–96, 2020.

Merci pour votre attention

Victoria Eyharabide

maria-victoria.eyharabide@sorbonne-universite.fr

Détection d'enluminures dans des manuscrits médiévaux IIF à l'aide de l'apprentissage profond

Victoria Eyharabide¹ Fouad Aouinti¹ Xavier Fresquet²

¹Laboratoire STIH, Sorbonne Université, France

²IReMus Laboratory, Sorbonne Université, France

24 mars 2021

Article soumis à ICDAR 2021 (16th International Conference on Document Analysis and Recognition)