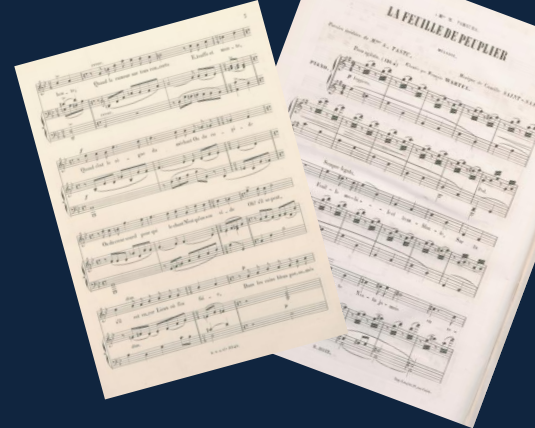




Combinaison d'IA pour la reconnaissance de partitions musicales

Aurélie Lemaitre – IRISA – Université Rennes 2

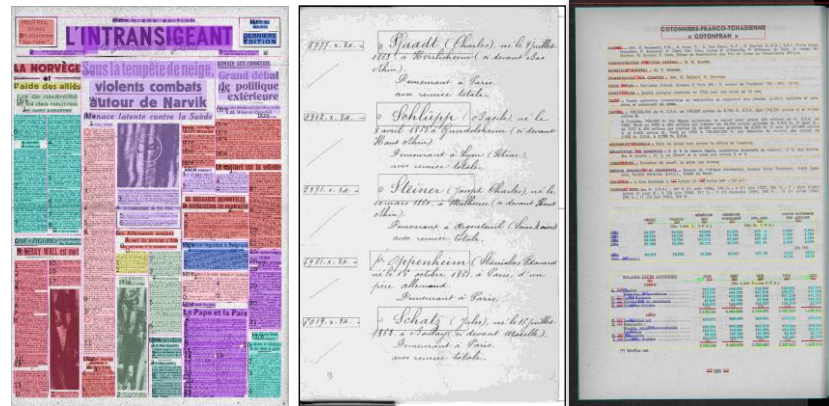
ARDOISE – 21 janvier 2026

Professeur en informatique

- Enseignement département MIASHS – Université Rennes 2
- Recherche IRISA – Equipe Shadoc

Systèmes Hybrides d'Analyse de DOcuments

- Analyses d'images de documents scannés
- Extraction de contenu



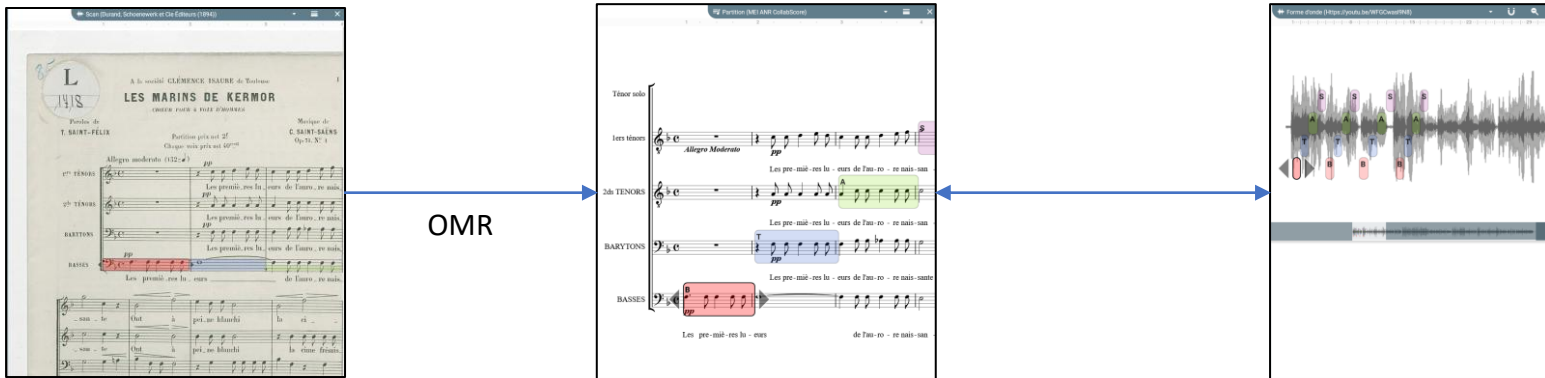
Projet ANR 2021 - 2025

Partenaires

- CNAM, BNF, Iremus, IRISA, Algomus, Fondation Royaumont

Objectifs

- Reconnaissance de partitions musicales
- Valorisation de fonds musicaux de la BNF
- Alignement entre image et enregistrement audio





Requête à ChatGPT (Open AI)

- Production d'un résultat qui n'a rien à voir



Requête à Le Chat (Mistral AI)

- Répond qu'il ne sait pas faire
- Conseille d'utiliser un logiciel d'OMR

Pourquoi ces limites ?

- Complexité de la structure musicale
 - Règles d'écritures en 2 dimensions
 - Faciles à expliquer, difficiles à inférer
 - Peu de données annotées disponibles
 - Entrée : image
 - Ex : IMSLP : bases de données de partitions numérisées (pdf...)
 - Sortie : fichier XML / Midi
 - Ex : Musescore.org : musique numérique (Xml)
- => Pas de lien entre les deux !



Applicables à des documents de bonne qualité

- Edition récente
- Format pdf

=> Difficulté sur des documents anciens

Bonne analyse des structures fréquentes

- Rythmes simples, partition monophoniques

=> Difficulté sur des partitions au contenu musical complexe

- Partitions polyphoniques, portées doubles (piano...)

=> Nécessite une relecture manuelle intégrale et fastidieuse

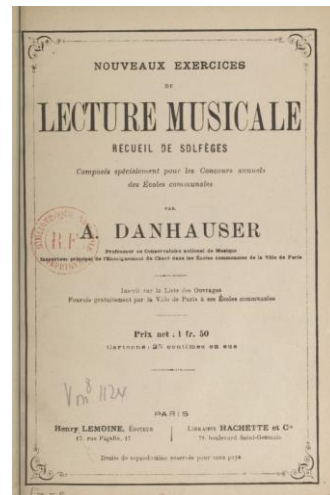
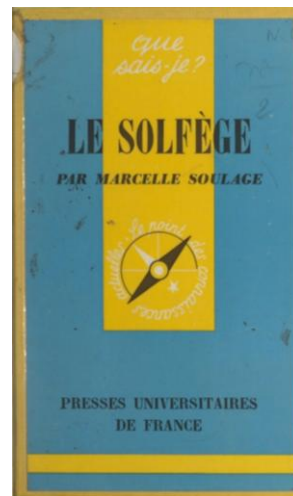


Apprentissage du solfège humain

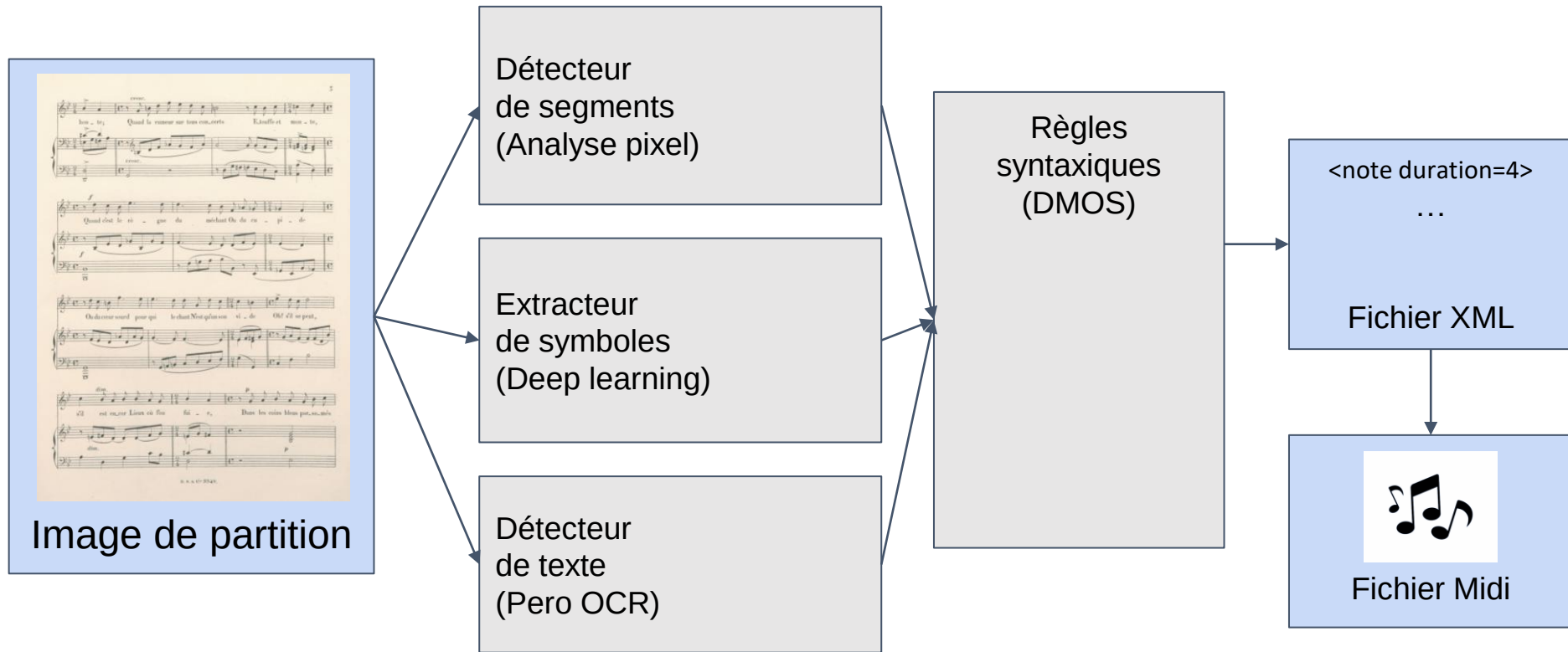
IA Hybride

- Exemples de symboles
- Règles de la notation musicale

=> Construction d'un système OMR hybride



Architecture générale : combinaison d'IA

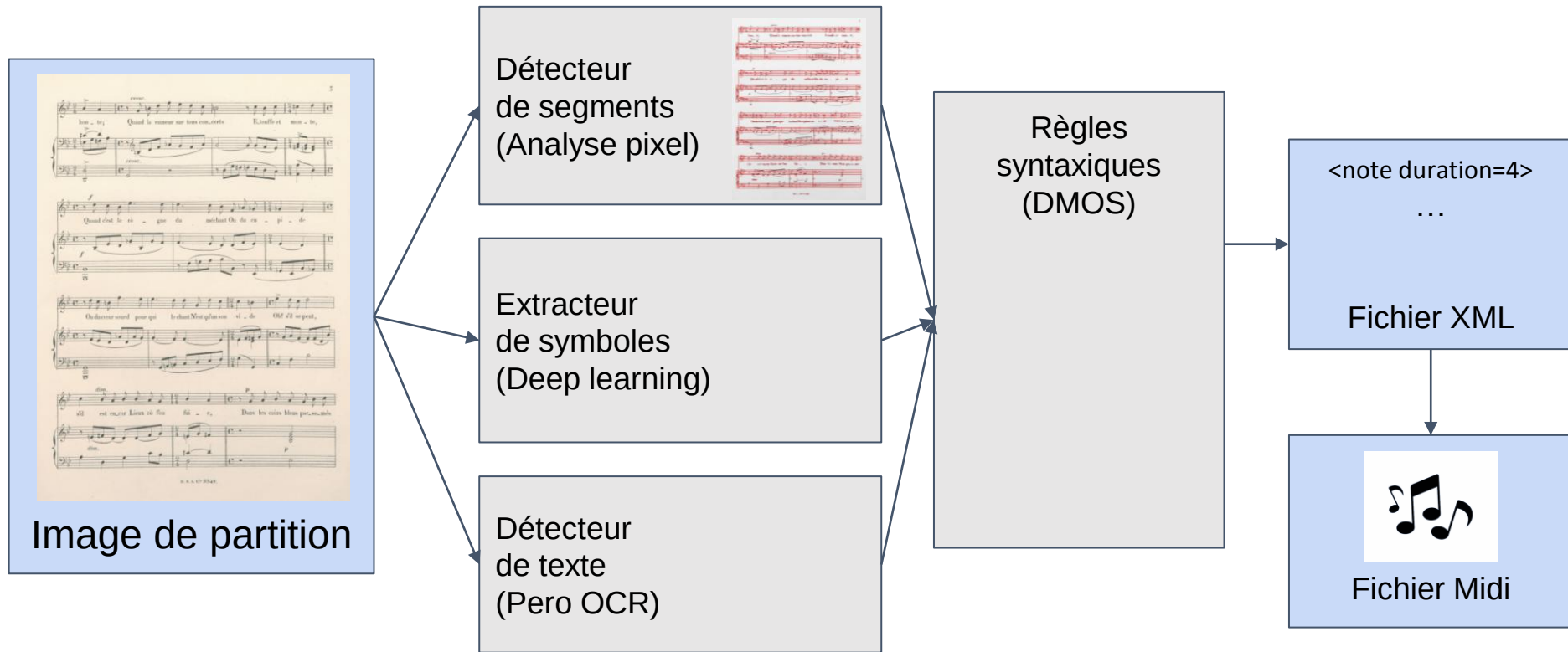


Analyse d'image « traditionnelle »

- Filtrage de Kalman
 - Analyse des pixels noirs qui se suivent
- Détection des portées, barres de mesures, hampes, barres de groupes



Architecture générale : combinaison d'IA



Objectif

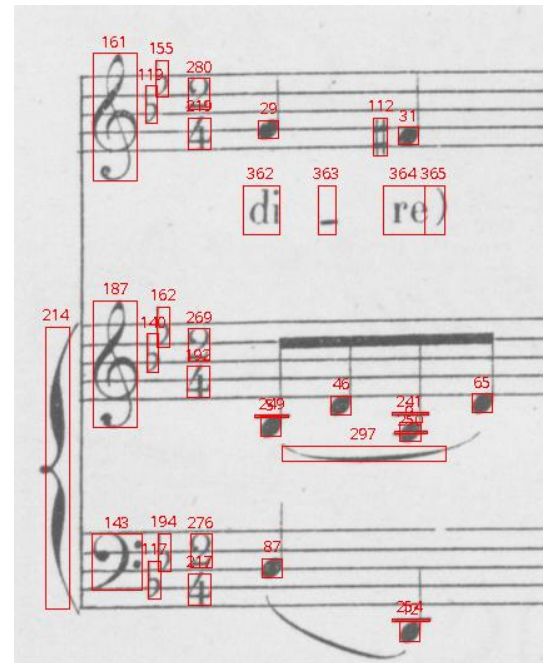
- Localisation des symboles musicaux dans une image
- Entraînement d'un réseau de neurones profonds

Données d'entraînement

- Besoin de données étiquetées : pixels avec label
- Peu de données disponibles sur les partitions anciennes

Spécificité de la détection

- Détection sur une page complète (3000x2000 pixels)
- Détection de tous petits symboles



Sélection d'une architecture Cascade R-CNN + Focal Net

Entraînement sur données synthétiques

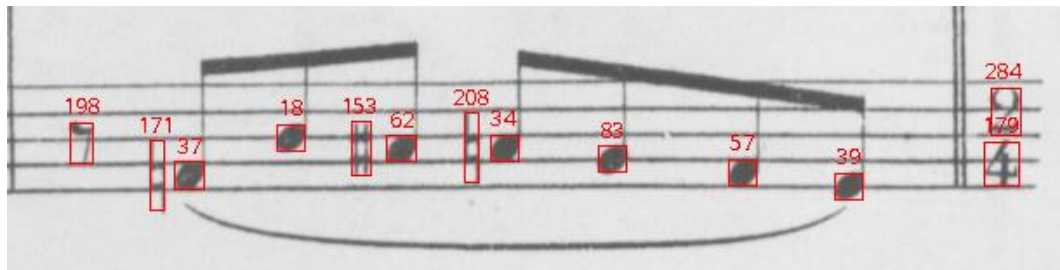
- Base "Deep Scores V2 Dense"
- 1714 images, 1.1M annotations, 136 classes

Fine-tuning sur données réelles

- Annotation de 2 bases du XIX^e siècle par bootstrap
- BNF-20-pages et BNF-25-pages

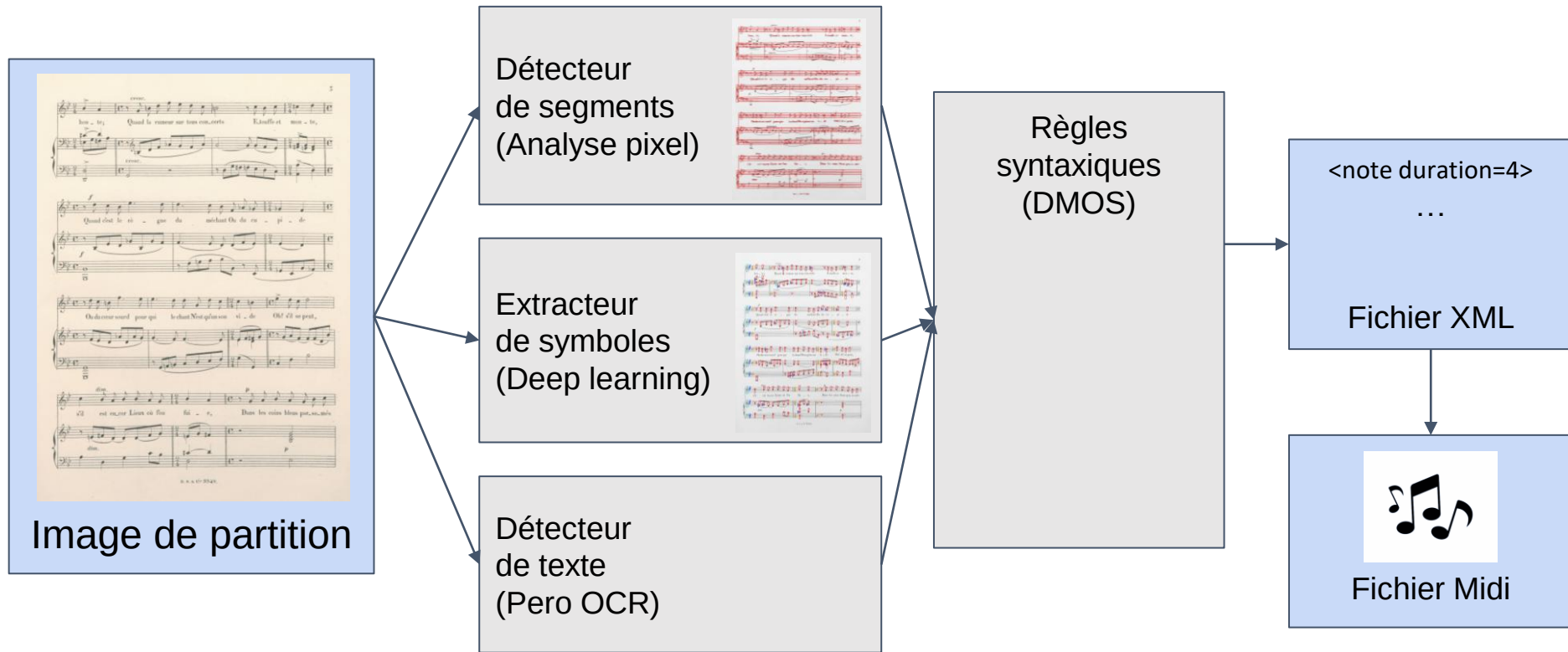
Stratégie pour une meilleure performance

- Limiter le nombre de symboles
 - 63 classes de symboles isolés à forme fixe
 - Clés, têtes de notes, chiffrage, altérations, nuances
- Ôter les symboles à forme variable



=> Rendu possible car combinaison d'IA

Architecture générale : combinaison d'IA



Utilisation d'un OCR open source

- Pero OCR

Détection du texte

- Paroles
- Indications d'interprétation

Décomposition en syllabes

CALME DES NUITS

C. SAINT-SAËNS
Op. 68, No. 1

Molto Adagio pp

SOPRANOS
Cal me des nuits, frai cheur des soirs,

CONTRALTOS
pp
Cal me des nuits, frai cheur des soirs,

TÉNORS
pp
Cal me des nuits, frai cheur des soirs,

BASSES
pp
Cal me des nuits, frai cheur des soirs,

PIANO
ad libitum pp

Molto Adagio

Architecture générale : combinaison d'IA

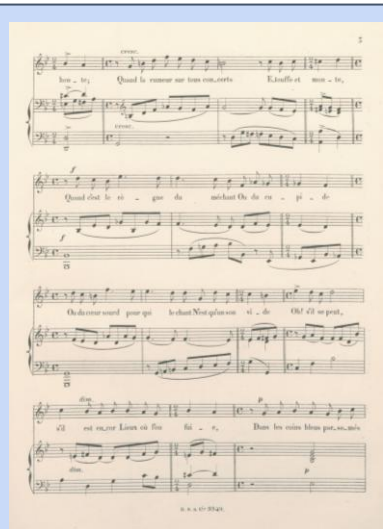
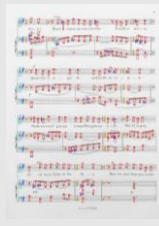


Image de partition

Détecteur
de segments
(Analyse pixel)



Extracteur
de symboles
(Deep learning)



Détecteur
de texte
(Pero OCR)



Modélisation
syntaxique
(DMOS)

<note duration=4>
...

Fichier XML



Fichier Midi

Objectif

- Combiner des données hétérogènes avec des règles

Utilisation de la méthode DMOS

- Méthode de reconnaissance de documents
- Définition de règles syntaxiques en langage EPF

Appliquée à des domaines différents

- Documents manuscrits/imprimés, hétérogènes
- Documents anciens, abîmés

Intérêt

- Décrire simplement les concepts musicaux complexes à reconnaître



```
partition :=
  AT(hautPage) &&
  portee P &&
  AT(surPortee P) &&
  symbolesMusicaux.
```

Connaissances de l'utilisateur

Règles de description
des connaissances

Compilation

Collection de
documents

Extraction
d'indices visuels

Analyseur dédié à
la collection

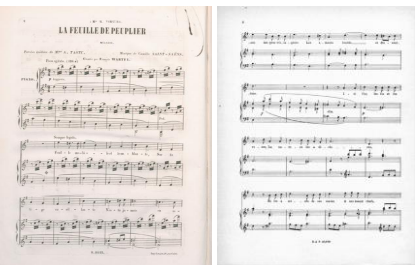
Résultat
d'analyse

Analyse image : portées

Réseaux de neurones :
symboles

OCR : paroles

Combinaison d'IA



porteeSimple :=

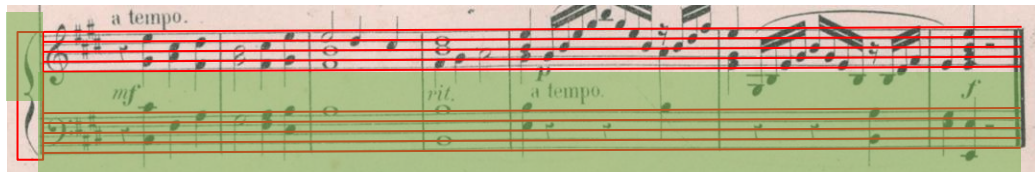
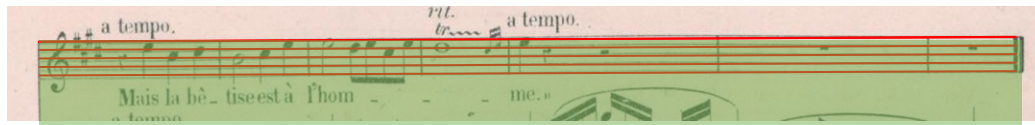
ligneHorizontale L &&
 AT(sousLigne L) &&
 quatreLignesHorizontales.

porteeDouble :=

porteeSimple P1 &&
 AT(gauchePortee P1) &&
 accolade A &&
 AT(sousPortee P1) &&
 porteeSimple P2.

systeme :=

suitePorteeSimpleOuDouble.



Page 19 of a musical score. The page is divided into two systems, each containing 12 staves. The top system is marked with a red bracket and labeled 'Système 1' and '12 portées simples'. The bottom system is marked with a red bracket and labeled 'Système 2' and '12 portées simples'. The score includes various musical notations such as notes, rests, and dynamic markings like 'pp' and 'f'.

Système 1
12 portées simples

Système 2
12 portées simples

Page 3 of a musical score. The page is divided into five systems, each containing 2 staves. The systems are labeled with red brackets and text: 'Système 1' (1 portée simple, 1 portée double), 'Système 2' (1 portée simple, 1 portée double), 'Système 3' (1 portée simple, 1 portée double), 'Système 4' (1 portée simple, 1 portée double), and 'Système 5' (1 portée simple, 1 portée double). The score includes various musical notations such as notes, rests, and dynamic markings like 'pp' and 'f'.

Système 1
1 portée simple
1 portée double

Système 2
1 portée simple
1 portée double

Système 3
1 portée simple
1 portée double

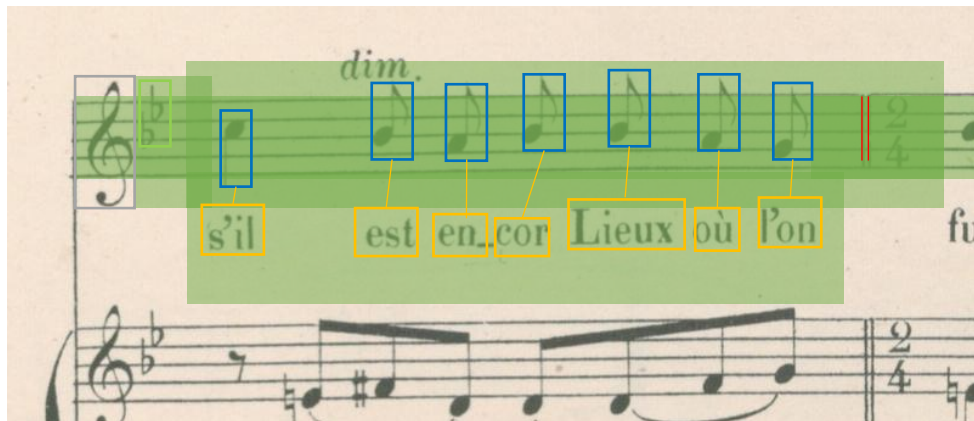
Système 4
1 portée simple
1 portée double

Système 5
1 portée simple
1 portée double

mesure :=

```

    AT(surPortee) &&
    clé C &&
    AT(droiteClé C) &&
    desAltérations A &&
    AT(droiteAltérations A) &&
    suiteNotesEtSilences N &&
    AT(sousNote N) &&
    parolesAssociées P N &&
    AT(droiteNotes N) &&
    barreFinMesure B.
  
```



Vocabulaire classique de la notation musicale

- Entête : clé, armure, chiffrage
- Notes : hauteur, rythme, altération, point
- Silences

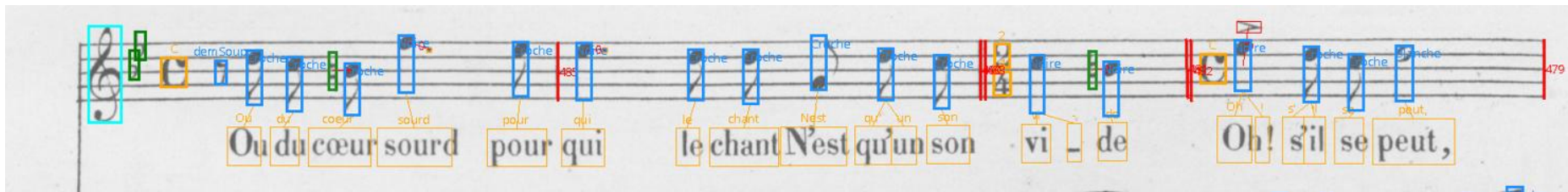
The image shows a musical score for three staves (treble, alto, and bass clef) with the lyrics "Quand c'est le règne du méchant Ou du cu-pi-de". The score is annotated with various graphical elements:

- Staff 1 (Treble Clef):** The first measure is highlighted with a red box. The second measure is highlighted with a green box. The third measure is highlighted with a blue box. The fourth measure is highlighted with a blue box. The fifth measure is highlighted with a blue box. The sixth measure is highlighted with a blue box. The seventh measure is highlighted with a blue box. The eighth measure is highlighted with a blue box. The ninth measure is highlighted with a blue box. The tenth measure is highlighted with a blue box. The eleventh measure is highlighted with a blue box. The twelfth measure is highlighted with a blue box. The thirteenth measure is highlighted with a blue box. The fourteenth measure is highlighted with a blue box. The fifteenth measure is highlighted with a blue box. The sixteenth measure is highlighted with a blue box. The seventeenth measure is highlighted with a blue box. The eighteenth measure is highlighted with a blue box. The nineteenth measure is highlighted with a blue box. The twentieth measure is highlighted with a blue box. The twenty-first measure is highlighted with a blue box. The twenty-second measure is highlighted with a blue box. The twenty-third measure is highlighted with a blue box. The twenty-fourth measure is highlighted with a blue box. The twenty-fifth measure is highlighted with a blue box. The twenty-sixth measure is highlighted with a blue box. The twenty-seventh measure is highlighted with a blue box. The twenty-eighth measure is highlighted with a blue box. The twenty-ninth measure is highlighted with a blue box. The thirtieth measure is highlighted with a blue box. The thirty-first measure is highlighted with a blue box. The thirty-second measure is highlighted with a blue box. The thirty-third measure is highlighted with a blue box. The thirty-fourth measure is highlighted with a blue box. The thirty-fifth measure is highlighted with a blue box. The thirty-sixth measure is highlighted with a blue box. The thirty-seventh measure is highlighted with a blue box. The thirty-eighth measure is highlighted with a blue box. The thirty-ninth measure is highlighted with a blue box. The fortieth measure is highlighted with a blue box. The forty-first measure is highlighted with a blue box. The forty-second measure is highlighted with a blue box. The forty-third measure is highlighted with a blue box. The forty-fourth measure is highlighted with a blue box. The forty-fifth measure is highlighted with a blue box. The forty-sixth measure is highlighted with a blue box. The forty-seventh measure is highlighted with a blue box. The forty-eighth measure is highlighted with a blue box. The forty-ninth measure is highlighted with a blue box. The fiftieth measure is highlighted with a blue box. The fifty-first measure is highlighted with a blue box. The fifty-second measure is highlighted with a blue box. The fifty-third measure is highlighted with a blue box. The fifty-fourth measure is highlighted with a blue box. The fifty-fifth measure is highlighted with a blue box. The fifty-sixth measure is highlighted with a blue box. The fifty-seventh measure is highlighted with a blue box. The fifty-eighth measure is highlighted with a blue box. The fifty-ninth measure is highlighted with a blue box. The sixtieth measure is highlighted with a blue box. The sixty-first measure is highlighted with a blue box. The sixty-second measure is highlighted with a blue box. The sixty-third measure is highlighted with a blue box. The sixty-fourth measure is highlighted with a blue box. The sixty-fifth measure is highlighted with a blue box. The sixty-sixth measure is highlighted with a blue box. The sixty-seventh measure is highlighted with a blue box. The sixty-eighth measure is highlighted with a blue box. The sixty-ninth measure is highlighted with a blue box. The seventieth measure is highlighted with a blue box. The seventy-first measure is highlighted with a blue box. The seventy-second measure is highlighted with a blue box. The seventy-third measure is highlighted with a blue box. The seventy-fourth measure is highlighted with a blue box. The seventy-fifth measure is highlighted with a blue box. The seventy-sixth measure is highlighted with a blue box. The seventy-seventh measure is highlighted with a blue box. The seventy-eighth measure is highlighted with a blue box. The seventy-ninth measure is highlighted with a blue box. The eightieth measure is highlighted with a blue box. The eighty-first measure is highlighted with a blue box. The eighty-second measure is highlighted with a blue box. The eighty-third measure is highlighted with a blue box. The eighty-fourth measure is highlighted with a blue box. The eighty-fifth measure is highlighted with a blue box. The eighty-sixth measure is highlighted with a blue box. The eighty-seventh measure is highlighted with a blue box. The eighty-eighth measure is highlighted with a blue box. The eighty-ninth measure is highlighted with a blue box. The ninetieth measure is highlighted with a blue box. The ninety-first measure is highlighted with a blue box. The ninety-second measure is highlighted with a blue box. The ninety-third measure is highlighted with a blue box. The ninety-fourth measure is highlighted with a blue box. The ninety-fifth measure is highlighted with a blue box. The ninety-sixth measure is highlighted with a blue box. The ninety-seventh measure is highlighted with a blue box. The ninety-eighth measure is highlighted with a blue box. The ninety-ninth measure is highlighted with a blue box. The hundredth measure is highlighted with a blue box.

Problème complexe peu traité dans la littérature

Description syntaxique simple

- Construction de lignes de paroles sous les portées simples
- Attribution de chaque syllabe à la note la plus proche
- Production de tirets pour les mélismes / continuation de mots



Gestion des couplets multiples

Diagram illustrating the management of multiple stanzas (couplets) in a musical score. The score is divided into three systems, each containing three staves (treble, alto, and bass clef). The lyrics are: "A no-tre Roi! Grand pour l'His-toi-re, God! save the King! Send him vic-tori-ous,". The diagram shows how the same musical notation is repeated across different systems, with the lyrics being shifted to match the corresponding stanzas. The first system is marked with a green box and the number 144. The second system is marked with a red box and the number 186. The third system is marked with a red box and the number 189. The lyrics are repeated across these systems, with the first system showing the full lyrics and the subsequent systems showing the lyrics shifted to match the corresponding stanzas.

Diagram illustrating the management of multiple stanzas (couplets) in a musical score. The score is divided into two systems, each containing two staves (treble and bass clef). The lyrics are: "A no-tre Roi! Grand pour l'His-toi-re, God! save the King! Send him vic-tori-ous,". The diagram shows how the same musical notation is repeated across different systems, with the lyrics being shifted to match the corresponding stanzas. The first system is marked with a green box and the number 5. The second system is marked with a red box and the number 186. The lyrics are repeated across these systems, with the first system showing the full lyrics and the subsequent systems showing the lyrics shifted to match the corresponding stanzas.

Règles musicales simples à décrire

- Nombre de temps attendu dans la mesure
- Propagé depuis le début de la partition

Synchronisation verticale des notes

- Construction d'un grille rythmique

Quand c'est le règne

4 temps

4 temps

4 temps

4 temps

The image shows a musical score for the phrase "Quand c'est le règne". It features four staves. The first staff has a treble clef and a 4/4 time signature. The second staff has a bass clef. The third and fourth staves have a treble clef. Red vertical lines are drawn across all staves at regular intervals, representing a 4-beat grid. Notes are marked with blue boxes and labeled "Noire" (quarter note). Some notes are marked with yellow boxes and labeled "C" (C-clef). The lyrics "Quand c'est le règne" are written below the staves.

pi de

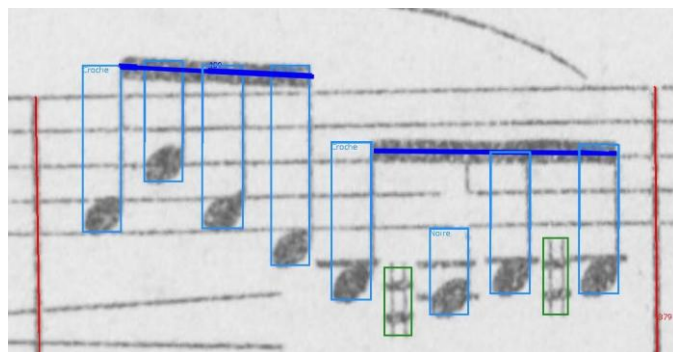
4 temps

4 temps

4 temps

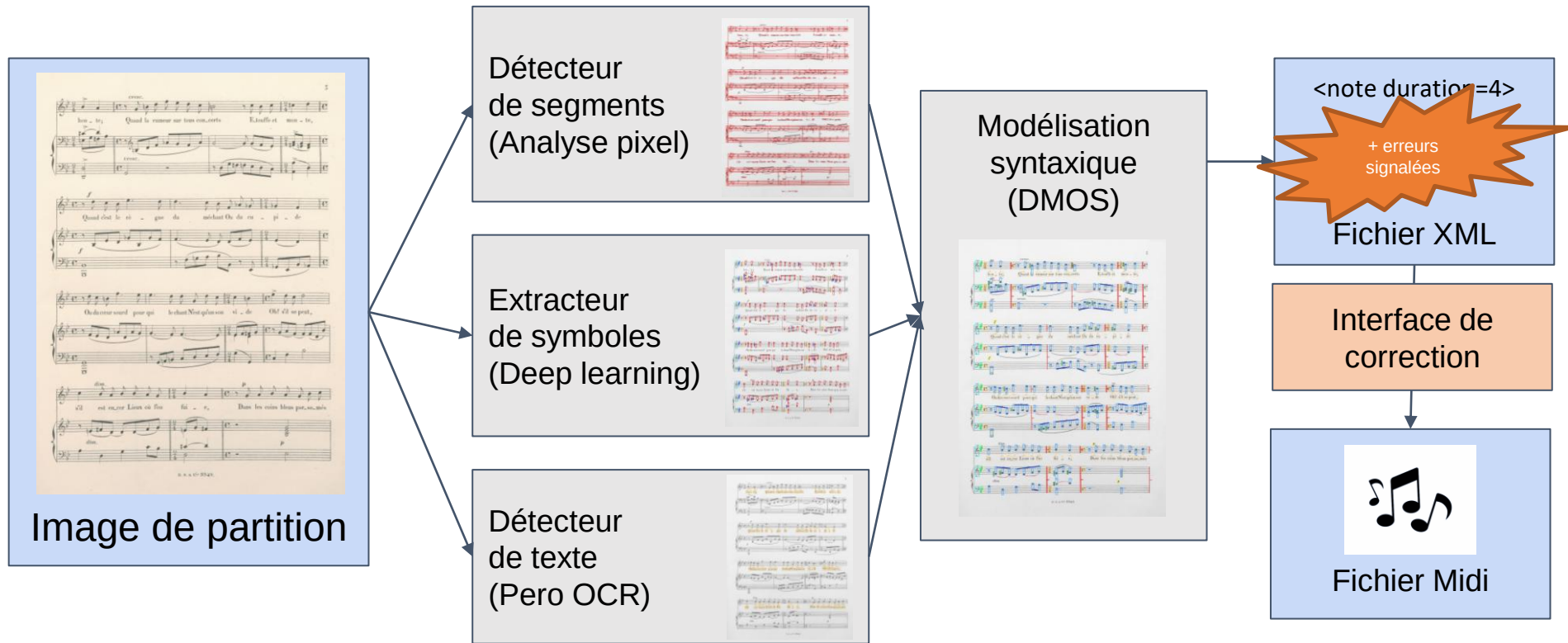
The image shows a musical score for the phrase "pi de". It features three staves. The first staff has a treble clef and a 4/4 time signature. The second staff has a bass clef. The third staff has a treble clef. Red vertical lines are drawn across all staves at regular intervals, representing a 4-beat grid. Notes are marked with blue boxes and labeled "Noire" (quarter note). Some notes are marked with yellow boxes and labeled "C" (C-clef). The lyrics "pi de" are written below the staves.

Détection d'incohérences temporelles



=> Vérification manuelle uniquement sur les erreurs signalées

Architecture générale : combinaison d'IA



Interface mise au point par le CNAM

- Lecture du résultat reconnu
- Mise en évidence des éléments à corriger par l'utilisateur

The screenshot displays the ColabScore web application interface. At the top, a navigation bar includes the ColabScore logo and various project links. The main content area shows a musical score for the piece "DANS LES COINS BLEUS" by C. SAINT-SAËNS. The score is presented in a multi-staff format, including vocal parts (Soprano, Alto, Tenor, Bass) and piano accompaniment. A circular correction tool is overlaid on the score, allowing users to interact with the notes. The right sidebar contains a "Tâches" (Tasks) section with a "Page 1/4" indicator and a "Valider cette page" (Validate this page) button. Below this, there are links for "Instructions phase 2", "Instructions phase 3", and "Explications détaillées". A section titled "Liste des erreurs trouvées dans cette page" (List of errors found on this page) indicates that no errors were detected on this page. The bottom of the interface shows the version number "Version 0.6.4" and the text "Construit par Teklia - Documentation".

Pages



1



2



3



4

2

DANS LES COINS BLEUS

Précédent de
SAINT-BEUVE

Suivant de
C. SAINT-SAËNS

Mod^o non troppo lento

CHANT

Dans les coins bleus par-se-més d'or (Sans trop le

PIANO

Mod^o non troppo lento

di - re) Il faut qu'on cache âme et tré - sor, Et doux mar -

ty - re. Quand tout dé-borde en l'a - ni - vers, Quand nul n'a

Durand, Schenker & Co Éditeurs. D. 8 n° 5349 Paris, 4, Place de la Madeleine.

Jouer ▶

Tâches

◀ Préc.

Page 1/4

Suiv. ▶

Phase 3 : Objets musicaux

Valider cette page

Instructions phase 3

Explications détaillées

Liste des erreurs trouvées dans cette page

Aucune erreur n'a été détectée sur cette page, mais rien de vous empêche de vérifier par vous-mêmes

https://dezrann.net/~collabscore/all:collabscore:concert:C144_0

Partition (page) (Partition imprimée (BnF))

0:00 0:05 0:10 0:15 0:20

à Madame GAVEAUX-SARATIER.

RÉVERIE

Paroles de V. HUGO. Musique de C. SAINT-SAËNS.

Allegretto tranquillo.

CHANT.

Puisqu'ici bas toute â - me Donne à quel

PIANO.

qu'un Sa mu-si-que sa flamme Ou son par-fum Puis-qu'ici toute

cho-se Donne toujours Son épine ou sa ro-se A ses a-mours Puisque faire à la

Partition (page) (Paged Score)

0:00 0:05 0:10 0:15

Chant

Puisqu'ici bas toute â - me Donne à quel

Piano

5

C.

u'un Sa mu-si-que sa flamme Ou son par-fum Puis-qu'ici toute

p.

MODIFIER

Type



Tag

Commentaire

+ AJOUTER

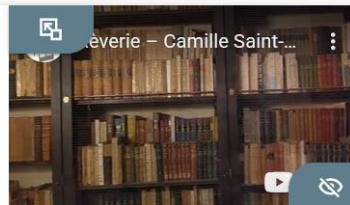
FILTRE

Début

Durée

Fin

✓ CACHER



Production d'un système OMR (Optical Music Recognition)

Combinaison d'IA

- Analyse d'image traditionnelle
- Entraînement d'un réseau de neurones
- Guidé par méthode de règles syntaxiques

Intérêt de la combinaison

- Bénéficier des capacités d'extracteurs de symboles
- Limiter les besoins en données annotées
- Exprimer facilement les règles logiques musicales
- Fiabiliser le résultat en pointant les erreurs à l'utilisateur

Prochains travaux

- Adaptation du système aux partitions manuscrites (Déecteur d'objets)
- Construction de systèmes de bout en bout entraînés avec l'OMR CollabScore

FOLLOW US!



www.irisa.fr



[irisa-lab](#)



Institut de Recherche en Informatique et Systèmes Aléatoires



Université
de Rennes



CentraleSupélec



école
normale
supérieure



IMT Atlantique
Département des Sciences
de l'Informatique



INSA

INSTITUT NATIONAL
DES SCIENCES
APPLIQUÉES
RENNES

Inserm



Peu de données annotées disponibles sur les partitions anciennes

Données synthétiques : base “Deep Scores V2 Dense”

- 136 Classes, 1714 Images, 1.1M annotations

Sélection de l'architecture Cascade R-CNN+FocalNet

- **Test sur 63 classes**

Model (Detector - Backbone)	#Classes in Train	mAP@0.5	Weighted mAP@0.5	mAP	Weighted mAP
Faster RCNN - HRNet	63	0.994	0.994	0.881	0.911
Cascade RCNN - HRNet	63	0.992	0.992	0.920	0.939
Cascade RCNN + FocalNet	63	0.996	0.992	0.929	0.940
Cascade RCNN + FocalNet	136	0.835	0.670	0.773	0.625



- **Dégradation si symboles à forme variable (136 classes)**
=> Intérêt de l'Hybridation Deep et Syntaxique

Données réelles du XIXe siècle : Fine-tuning + Augmentation

- Pages réelles annotées par bootstrap avec modèle sur données synthétiques (Deep Score)
 - 25 pages : BNF-25, 20 pages BNF-20
- Pages réelles pseudo-labels avec modèle précédent
 - 197 pages : BNF-197

	Evaluation on BNF-25				Evaluation on BNF-20			
Cascade RCNN - FocalNet (Proposed) / 63 Classes	mAP@0.5	Weighted mAP@0.5	mAP	Weighted mAP	mAP@0.5	Weighted mAP@0.5	mAP	Weighted mAP
Deep Scores V2	0.776	0.974	0.755	0.964	0.771	0.973	0.753	0.963
Deep Scores V2 + BNF 20 + BNF-197-20	0.870	0.987	0.800	0.936	-	-	-	-
Deep Scores V2 + BNF 25 + BNF-197-25	-	-	-	-	0.833	0.991	0.752	0.940