

UTILISATION DU TALN POUR L'ANALYSE DE DOCUMENTS D'ARCHIVES MANUSCRITES : APPLICATION AUX REGISTRES DU THÉÂTRE DE LA COMÉDIE-ITALIENNE

MeetUp Machine Learning

Adeline GRANET

5 décembre 2016

LINA équipe TALN - IRCCyN équipe IVC
Université de Nantes

PROJET CIRESFI

Chronologie de l'Opéra-Comique

1669 Louis XIV fonde l'Académie Royale de musique (Opéra).

1680 Louis XIV fonde la Comédie-Française et leur donne le *privège* d'être les seuls lieux de représentations de la musique et du théâtre.

1697 Les comédiens italiens n'ont plus le droit de jouer et sont renvoyés de Paris.

1716 Le régent (Duc d'Orléans) autorise de nouveau les comédiens italiens à jouer à Paris.

⇒ **Début des registres étudiés.**

1762 Fusion de l'Opéra et de la Comédie-Italienne

1780 Interdiction de jouer en italien, la troupe reprend le nom d'Opéra-Comique.

1783 Déménagement dans le nouveau théâtre *Salle Favard* avec plus de 1100 places (aujourd'hui le *théâtre national de l'Opéra-Comique*)

1791-1793 Fin des privilèges et liberté des théâtres est proclamée ⇒ **Fin des registres étudiés**

Contrainte et Intégration : pour une Réévaluation des Spectacles Forains et Italiens sous l'Ancien Régime

1. **Contextualisation historico-politique** : intégration jusqu'à l'obtention du privilège de l'Opéra ;

Contrainte et Intégration : pour une Réévaluation des Spectacles Forains et Italiens sous l'Ancien Régime

1. **Contextualisation historico-politique** : intégration jusqu'à l'obtention du privilège de l'Opéra ;
2. **La contrainte comme moteur** : révolution dans la forme du théâtre avec des nouveaux genres de pièces par écriteaux, monologues à plusieurs, pantomimes, opéra-comique ;

Contrainte et Intégration : pour une Réévaluation des Spectacles Forains et Italiens sous l'Ancien Régime

1. **Contextualisation historico-politique** : intégration jusqu'à l'obtention du privilège de l'Opéra ;
2. **La contrainte comme moteur** : révolution dans la forme du théâtre avec des nouveaux genres de pièces par écriteaux, monologues à plusieurs, pantomimes, opéra-comique ;
3. **L'économie des spectacles** : retracer les coûts de production et organisation ;

Contrainte et Intégration : pour une Réévaluation des Spectacles Forains et Italiens sous l'Ancien Régime

1. **Contextualisation historico-politique** : intégration jusqu'à l'obtention du privilège de l'Opéra ;
2. **La contrainte comme moteur** : révolution dans la forme du théâtre avec des nouveaux genres de pièces par écriteaux, monologues à plusieurs, pantomimes, opéra-comique ;
3. **L'économie des spectacles** : retracer les coûts de production et organisation ;
4. **Exploration et analyse innovante** : numérisation, extraction d'information, et indexation des informations contenus dans les ressources disponibles.

Contrainte et Intégration : pour une Réévaluation des Spectacles Forains et Italiens sous l'Ancien Régime

1. **Contextualisation historico-politique** : intégration jusqu'à l'obtention du privilège de l'Opéra ;
 2. **La contrainte comme moteur** : révolution dans la forme du théâtre avec des nouveaux genres de pièces par écriteaux, monologues à plusieurs, pantomimes, opéra-comique ;
 3. **L'économie des spectacles** : retracer les coûts de production et organisation ;
 4. **Exploration et analyse innovante** : numérisation, extraction d'information, et indexation des informations contenus dans les ressources disponibles.
- ⇒ Défi pluridisciplinaire : mettre la reconnaissance de l'écriture et le traitement du langage au service de l'Histoire

Projet CIRESEFI

Les registres de la Comédie-Italienne

État de l'art : la reconnaissance de l'écriture

- Prétraitement

- Extraction des caractéristiques

- Méthodes de reconnaissance de l'écriture

- Traitement du langage

Identification des problèmes

Corpus et caractéristiques

- Les corpus

- Les features

Premiers tests et résultats

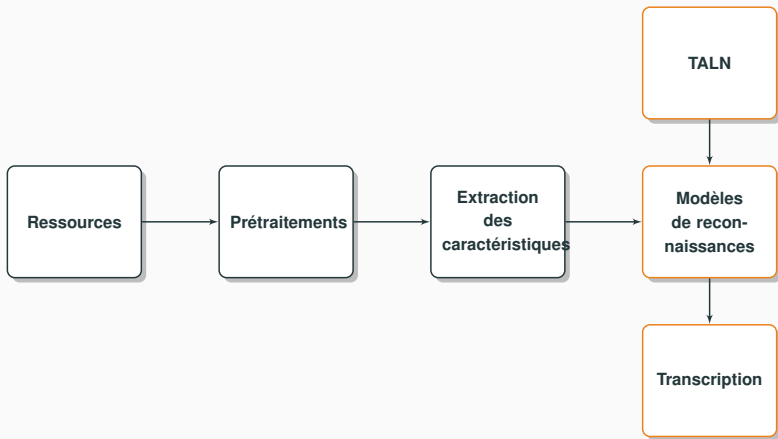


FIGURE – Présentation du processus complet dans le cadre de la reconnaissance d'écritures

LES REGISTRES DE LA COMÉDIE-ITALIENNE

ÉVOLUTION DES DOCUMENTS

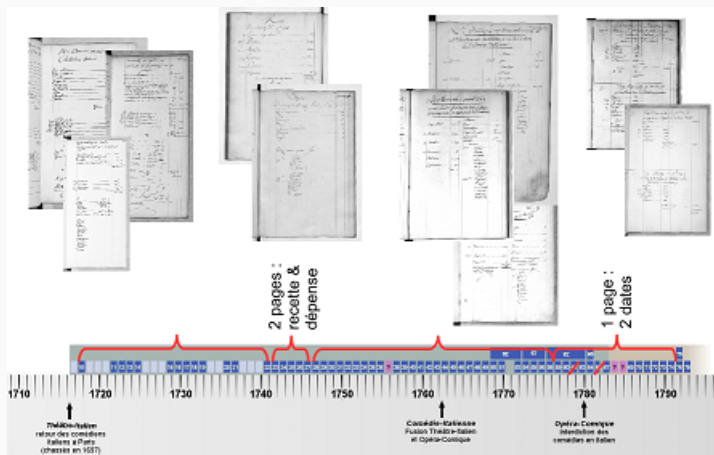


FIGURE – Frise reprenant l'ensemble des 63 registres par année et faisant mention de la forme utilisée sur la période 1717 à 1791.

ENSEMBLE DES DONNÉES À EXTRAIRE

Date

[illegible]

ENSEMBLE DES DONNÉES À EXTRAIRE

Date

Titre

16		Le Samedi 4 Juin 1768	
		<i>Première de Sophie ou Le Mariage caché. Comédie en trois actes mêlée d'ariettes. Suivie du Ballet Précedée d'opéra-ballet.</i>	
	Loges	6 22	Minos. - - - - - 2
			Jeux - - - - - 29 10
			Domestiques - - - - - 22
125	Premiers	250	page a la gausse - - - - - 23 10
			Amusements - - - - - 48
84	Secondes	252	Ballet - - - - - 1 10
			Manes
70	Troisiemes	140	Jeux Avant
			Jeux
			Jeux
	Quatriemes	4 12	Jeux
			Jeux
	Suppléments	1 16	Jeux
		2322	Jeux
	crédit		Jeux
	pour Mr Robault pour		Jeux
	22 Billets de 100 livres	72	Jeux
	pour Mr de la Roche pour		Jeux
	22 Billets de 100 livres	15	Jeux
	pour Mr de la Roche pour		Jeux
	22 Billets de 100 livres	74	Jeux

ENSEMBLE DES DONNÉES À EXTRAIRE

Date

Titre

Dépenses

Le Samedi 4 Juin 1768			
Premier de Sophie ou Le Mariage cache.			
Comédie en trois actes mêlée d'ariettes			
Suivie du Bal, & précédée d'arlog: toujours arloguée			
		Chœurs - - -	2
Loges - - - - -	6 25	Scène - - - - -	39 10
		Domestiques - - -	22
125 Parterres - - - -	7 50	Pages & laquais - -	23 10
		Luminières - - -	48
84 Secondes - - - -	2 52	Ballet - - - - -	1 10
		<u>Musique</u>	
70 Tiersiens - - - - -	1 40	Solles	favorit
			romant
			Scapins
			Amante
Parterres - - - - -	4 12		Compte
Suppléments - - - -	1 16		
	23 33	1 ^{re}	Comédie
crédit			Carlin
autres de la	pas m. Lehaud pour		Ballet
manque	52 Billets de Parterres - -	72	changé
Loges & par	pas pour coralline & me		Calot
supplément de	celle Loge de 2 ^e Comédie	15	Laruelle
des brats	autres		Cherrie
	pas m. Le Comte de Dumas		Dufrenoy
	pas de 1 ^{re} Loge de 2 ^e Com	74	trial

ENSEMBLE DES DONNÉES À EXTRAIRE

Date

Titre

Dépenses

Recettes

16		Le Samedi 4 Juin 1768		
		Projet de Sophie ou Le Mariage cache. Comedie en trois actes mêlée d'ariettes. Suivie du Ballet precedee d'arlog: (ou pour arlogues)		
		Loges	6 2 2	Minos. - - - - - 2
				Jeux - - - - - 29 10
125		Premieres	7 5 0	Domestiques - - - - - 22
				Pages a la garde - - - - - 23 10
84		Secondes	2 5 2	Amusements - - - - - 48
				Ballet - - - - - 1 10
70		Troisiemes	1 4 0	Manes
				Belle Avant
				Princo
				England
		Parties	4 12	Arquette
				Beaupe
		Supplements	1 6	
			23 25	pro - ciardi
				carlin
				Ballet
				Champille
				Calot
				Laruelle
				Clarice
				Defforms
				Trial

ENSEMBLE DES DONNÉES À EXTRAIRE

Date

Titre

Dépenses

Recettes

Acteurs

[illegible]

ENSEMBLE DES DONNÉES À EXTRAIRE

Date _____

Titre

Dépenses

Recettes

Acteurs

Notes

[illegible]

ENSEMBLE DES DONNÉES À EXTRAIRE

Date _____

Titre

Dépenses

Recettes

Acteurs

Notes

[illegible]

Objectifs

Étiqueter et indexer automatiquement chaque champs des registres journaliers

Extraire les entités nommées des registres

Divers types de documents : comptes journaliers, mensuels, annuels, état des pensionnaires, note volante, . . . ;

Mise en pages variée : plusieurs jours sur une page, deux pages pour une journée, une page pour une journée ;

Différentes langues : italien, vénicien, français ;

Plusieurs scripteurs : varie selon les registres ;

Disposition et informations changeantes : mention des jours de relâche, de fermeture du théâtre, précision sur la pièce jouée dans le titre,

Études précédentes

Temps : 3 mois pour chercher 1 information (droit d'auteurs) dans chaque compte mensuel sur une période de 10 ans des registres ;

Fiabilité : plusieurs transcriptions manuelles incomplètes et avec des interprétations.

Vocabulaire des titres basé sur *Repertorio*

de titres : 1055 ou 1182 (avec tous les titres d'une unique pièce)

de mots : 1359 (avec mots vides)

Les plus fréquents : Arlequin (197) (Arlequins(11), Arlequine(1), Arlequines(2)) ; amour (47) ; deux (31) ; Pantalon (23) ; amant (22) ; mariage (20) ; Scapin (16) ; amoureux (16) ;

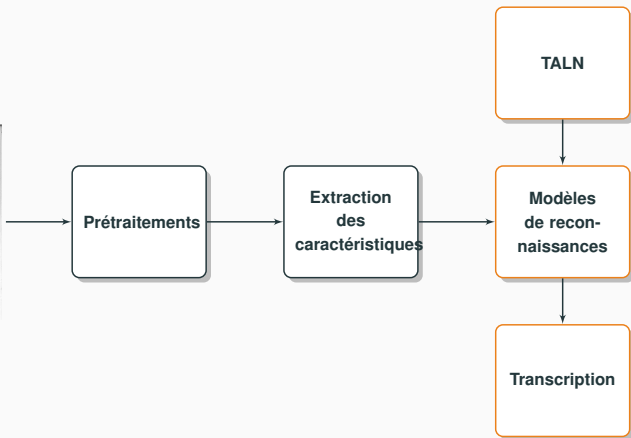
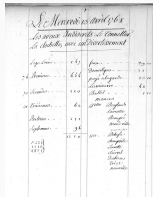
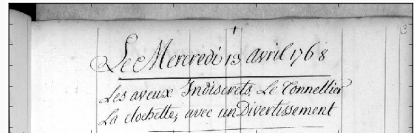


FIGURE – Présentation du processus complet dans le cadre de la reconnaissance d'écritures

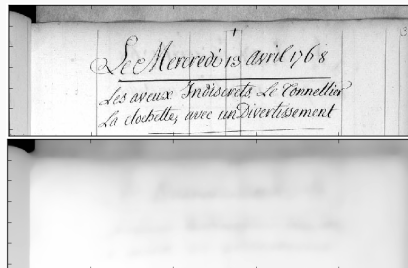
ÉTAT DE L'ART : LA RECONNAISSANCE DE L'ÉCRITURE

Conversion de l'image couleur
RGB en niveaux de gris



Conversion de l'image couleur
RGB en niveaux de gris

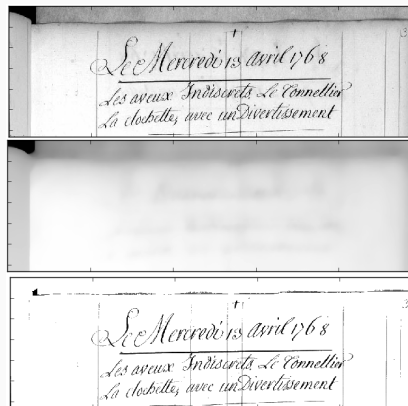
Suppression du fond de
l'image par filtre gaussien



Conversion de l'image couleur
RGB en niveaux de gris

Suppression du fond de
l'image par filtre gaussien

Binarisation de l'image avec
la méthode de Otsu



Conversion de l'image couleur
RGB en niveaux de gris

Suppression du fond de
l'image par filtre gaussien

Binarisation de l'image avec
la méthode de Otsu

Classification des pixels pour
extraire l'encre par CAE

Conversion de l'image couleur

RGB en niveaux de gris

Suppression du fond de
l'image par filtre gaussien

Binarisation de l'image avec
la méthode de Otsu

Classification des pixels pour
extraire l'encre par CAE

Limites

⇒ Risques de perdre des pixels
d'encre fondus dans le fond ;

⇒ Contour du document dans
l'image suivant la couleur peut
influencer les différentes
techniques

Les types de segmentation : bloc, ligne, mot

Les types de segmentation : bloc, ligne, mot

Les types de segmentation : bloc, ligne, mot

Segmentation en ligne par projection des pixels

Les types de segmentation : bloc, **ligne**, mot

Segmentation en ligne par projection des pixels

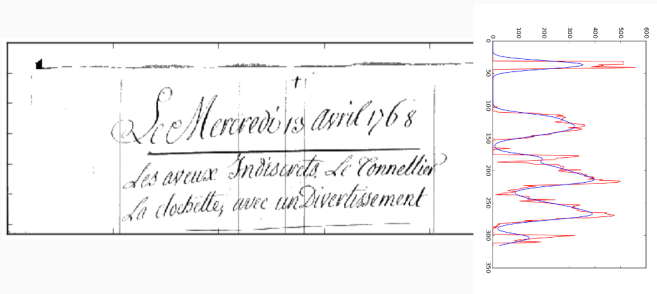


FIGURE – Projection de pixels (rouge), projection avec lissage (bleu)

Les types de segmentation : bloc, ligne, mot

Segmentation en ligne par projection des pixels

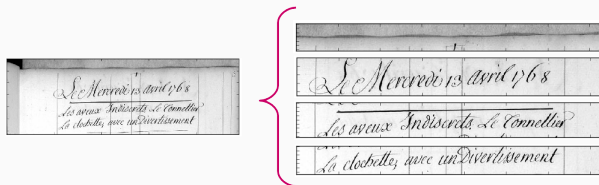


FIGURE – Segmentation par projection des pixels

Les types de segmentation : bloc, ligne, mot

Segmentation en ligne par projection des pixels

Regroupement de composantes connexes (avec une extraction de points d'intérêts, avec des projections de pixels et des HMMs, en utilisant la topologie des documents)

Les types de segmentation : bloc, ligne, mot

Segmentation en ligne par projection des pixels

Regroupement de composantes connexes (avec une extraction de points d'intérêts, avec des projections de pixels et des HMMs, en utilisant la topologie des documents)

Limites

- ⇒ Croisement entre les hampes et jambages de lignes connexes ;
- ⇒ Présence de filets potentiellement.
- ⇒ Solution DMOS envisagée

⇒ Diminuer les caractéristiques liées aux différents scripteurs

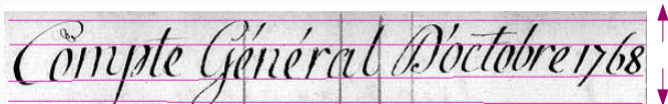


FIGURE – Normalisation par identification des lignes de base et fixation des espacements

Limites

⇒ Écriture cursive et longiligne plus prononcée à certaines périodes

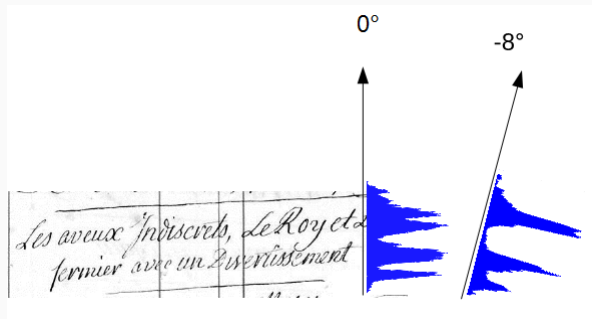


FIGURE – Correction de l'angle d'inclinaison des lignes par maximisation des projections de pixels

Limites

- ⇒ Écriture cursive et longiligne plus prononcée à certaines périodes
- ⇒ Beaucoup de lignes courbes dans les registres



FIGURE – Présentation du processus complet dans le cadre de la reconnaissance d'écritures

Caractéristiques géométriques et statistiques

Nombre de transition observée entre l'écriture et l'arrière-plan ;

Nombre de pixels d'encre observés ;

Position des contours supérieurs et inférieurs dans la fenêtre ;

Moyenne des valeurs des pixels ;

Position du centre de gravité ;

Position des lignes bases (ou références).

Caractéristiques directionnelles

SIFT : détection de points clés ;

HOG : histogrammes des gradients orientés.

Caractéristiques par AE, CAE et CNN

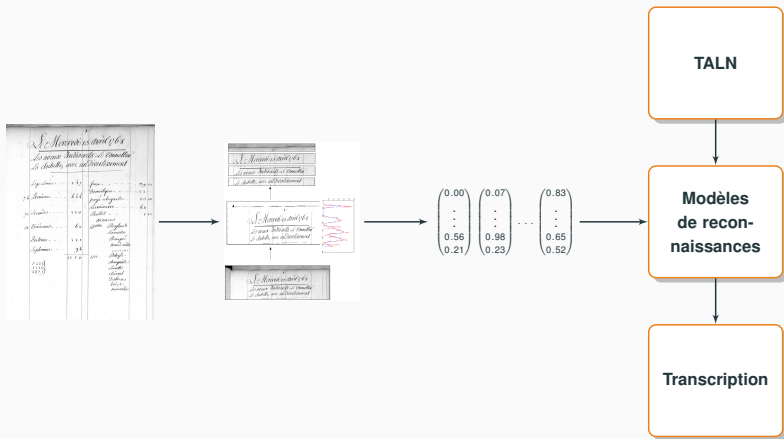


FIGURE – Présentation du processus complet dans le cadre de la reconnaissance d'écritures

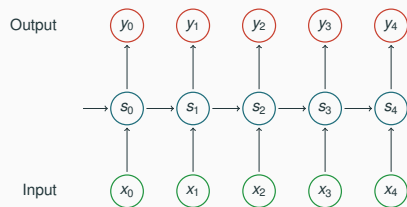


FIGURE – Réseau de neurones récurrents uni-directionnel

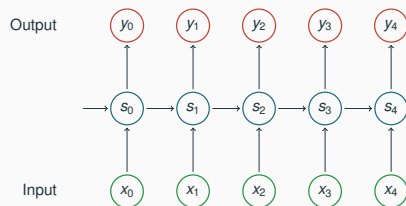


FIGURE – Réseau de neurones récurrents uni-directionnel

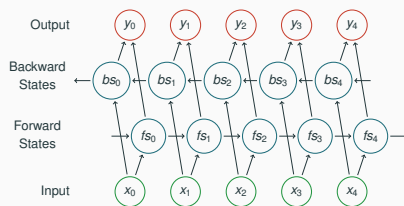


FIGURE – Réseau de neurones récurrents bi-directionnel

Limites

⇒ *Vanishing gradient* : perte du gradient sur le long terme

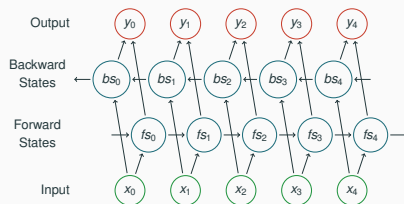


FIGURE – Réseau de neurones récurrents bi-directionnel

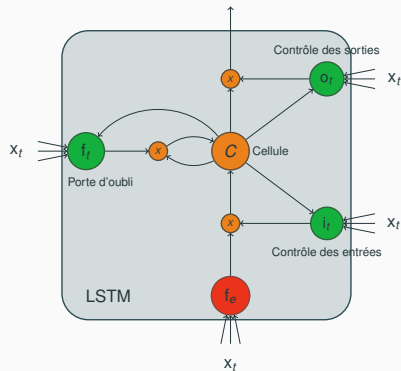


FIGURE – Cellule de type *Long short-term memory*

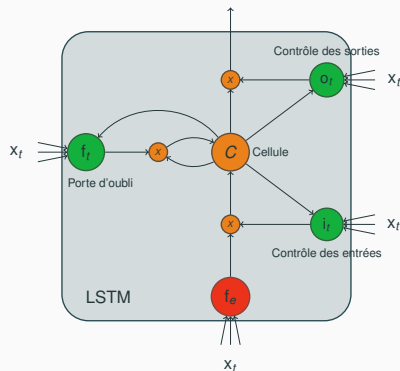


FIGURE – Cellule de type *Long short-term memory*

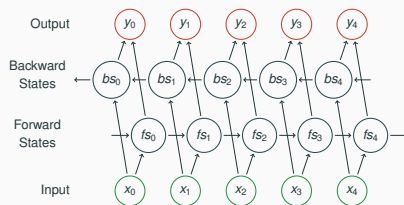


FIGURE – Réseau de neurones récurrents bidirectionnel utilisant les cellules LSTM

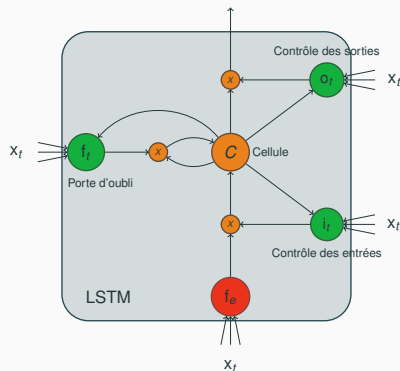


FIGURE – Cellule de type *Long short-term memory*

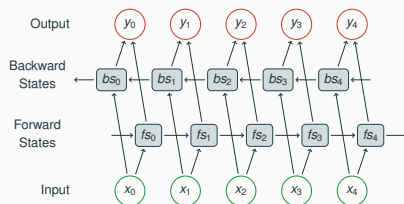


FIGURE – Réseau de neurones récurrents bidirectionnel utilisant les cellules LSTM

Créé pour des tâches de classifications temporels \Rightarrow problèmes d'étiquetages de séquences où l'alignement est inconnu entre l'entrée et la sortie du système

Pré-traitement des données d'entrée ou post-traitement des données en sortie du réseau : inutile

Prédit une séquence de caractères π pour une séquence d'entrée noté x

$$\mathbb{P}(\pi|x) = \prod_{t=1}^T y_{\pi_t}^t \quad (1)$$

où $y_{\pi_t}^t$ est la probabilité d'observer l'étiquette π à l'instant t

Caractéristiques

- Unique couche de neurones de type *Softmax*

- $|L|$ neurones où un neurone correspond à une classe

- +1 neurone "joker" appelé *blank*

Caractéristiques

Unique couche de neurones de type *Softmax*

$|L|$ neurones où un neurone correspond à une classe

+1 neurone "joker" appelé *blank*

Apprentissage : *algorithme forward-backward* modifié pour intégrer le label *blank* dans la séquence de label attendue

Apprentissage avec *algorithme forward-backward* modifié

Insertion du *blank* au début, à la fin et entre chaque caractère d'un label :

cat \Rightarrow *bl* c *bl* a *bl* t *bl*

$$|l'| = 2 * |l| + 1$$

Apprentissage avec *algorithme forward-backward* modifié

Insertion du *blank* au début, à la fin et entre chaque caractère d'un label :

cat \Rightarrow *bl c bl a bl t bl*

$$|l'| = 2 * || + 1$$

Calcul des variables :

$$\alpha_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{1:t} = l_{1:s})}} \prod_{t'=1}^t y_{\pi_{t'}}^{t'}$$

$$\beta_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{t:T} = l_{s:|l|})}} \prod_{t'=t}^T y_{\pi_{t'}}^{t'}$$

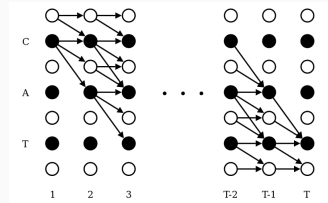


FIGURE — Treillis d'apprentissage pour le mot "cat" extrait de l'article de Graves et al

Apprentissage avec *algorithme forward-backward modifié*

Insertion du *blank* au début, à la fin et entre chaque caractère d'un label :

cat \Rightarrow *bl* c *bl* a *bl* t *bl*

$$|l'| = 2 * || + 1$$

Calcul des variables :

$$\alpha_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{1:t}=l_{1:s})}} \prod_{t'=1}^t y_{\pi_{t'}}^{t'}$$

$$\beta_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{t:T}=l_{s:|l|})}} \prod_{t'=t}^T y_{\pi_{t'}}^{t'}$$

Apprentissage avec *algorithme forward-backward* modifié

Insertion du *blank* au début, à la fin et entre chaque caractère d'un label :

cat \Rightarrow *bl* c *bl* a *bl* t *bl*

$$|l'| = 2 * |l| + 1$$

Calcul des variables :

$$\alpha_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{1:t}=l_{1:s})}} \prod_{t'=1}^t y_{\pi_{t'}}^{t'}$$

$$\beta_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{t:T}=l_{s:|l|})}} \prod_{t'=t}^T y_{\pi_{t'}}^{t'}$$

Fonction de perte et d'apprentissage :

Apprentissage avec *algorithme forward-backward* modifié

Insertion du *blank* au début, à la fin et entre chaque caractère d'un label :

cat \Rightarrow *bl c bl a bl t bl*

$$|l'| = 2 * |l| + 1$$

Calcul des variables :

$$\alpha_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{1:t}=l_{1:s})}} \prod_{t'=1}^t y_{\pi_{t'}}^{t'}$$

$$\beta_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{t:T}=l_{s:|l|})}} \prod_{t'=t}^T y_{\pi_{t'}}^{t'}$$

Fonction de perte et d'apprentissage :

$$O = - \sum \log(\mathbb{P}(l|x))$$

Apprentissage avec *algorithme forward-backward modifié*

Insertion du *blank* au début, à la fin et entre chaque caractère d'un label :

cat \Rightarrow *bl c bl a bl t bl*

$$|l'| = 2 * || + 1$$

Calcul des variables :

$$\alpha_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_1:t=l_1:s)}} \prod_{t'=1}^t y_{\pi_{t'}}^{t'}$$

$$\beta_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_t:T=l_s:|l|)}} \prod_{t'=t}^T y_{\pi_{t'}}^{t'}$$

Fonction de perte et d'apprentissage :

$$O = - \sum \log(\mathbb{P}(l|x))$$

$$\alpha_t(s)\beta_t(s) = \sum_{\substack{\pi \in \mathbb{B}^{-1}(l): t=1 \\ \pi_t=l'_s}} \prod_{t=1}^t y_{\pi_t}^t = \sum_{\substack{\pi \in \mathbb{B}^{-1}(l): \\ \pi_t=l'_s}} p(\pi|x)$$

Apprentissage avec *algorithme forward-backward modifié*

Insertion du *blank* au début, à la fin et entre chaque caractère d'un label :

cat \Rightarrow *bl c bl a bl t bl*

$$|l'| = 2 * |l| + 1$$

Calcul des variables :

$$\alpha_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{1:t}=l_{1:s})}} \prod_{t'=1}^t y_{\pi_{t'}}^{t'}$$

$$\beta_t(s) = \sum_{\substack{\pi \in \mathbb{N}^t: \\ \mathbb{B}(\pi_{t:T}=l_{s:|l|})}} \prod_{t'=t}^T y_{\pi_{t'}}^{t'}$$

Fonction de perte et d'apprentissage :

$$O = - \sum \log(\mathbb{P}(l|x))$$

$$\alpha_t(s)\beta_t(s) = \sum_{\substack{\pi \in \mathbb{B}^{-1}(l): \\ \pi_t=l'_s}} \prod_{t'=1}^t y_{\pi_{t'}}^{t'} = \sum_{\substack{\pi \in \mathbb{B}^{-1}(l): \\ \pi_t=l'_s}} p(\pi|x)$$

$$\mathbb{P}(l|x) = \sum_{s=1}^{|l'|} \alpha_t(s)\beta_t(s)$$

Caractéristiques

Unique couche de neurones de type *Softmax*

$|L|$ neurones où un neurone correspond à une classe

+1 neurone "joker" appelé *blank*

Apprentissage : *algorithme forward-backward* modifié pour intégrer le label *blank* dans la séquence de label attendue, suivie du *Maximum de vraisemblance* pour permettre la *backpropagation* du gradient

Décodage : *Meilleur chemin* qui est la concaténation à chaque instant du neurone le plus actif ; *Recherche de préfixe*

CONNEXIONIST TEMPORAL CLASSIFICATION [1] III

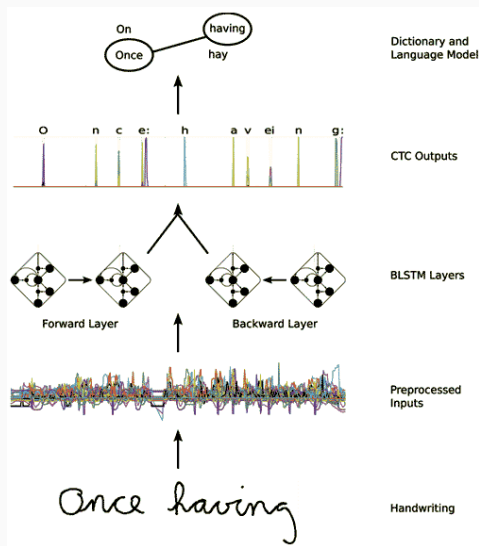


FIGURE – Système complet de reconnaissance d'écritures avec un réseau BLSTM et CTC extrait de l'article de Graves et *al*

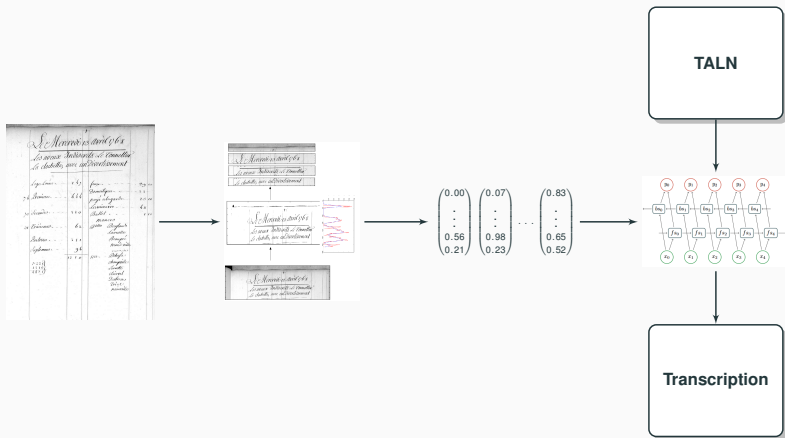


FIGURE – Présentation du processus complet dans le cadre de la reconnaissance d'écritures

Traitement du langage

Dictionnaire

Lexique : mot associé à sa forme de base

Méthode probabiliste des n-grammes

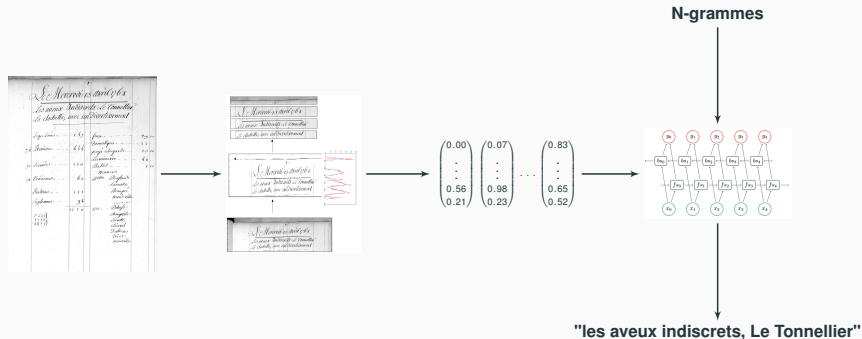


FIGURE – Présentation du processus complet dans le cadre de la reconnaissance d'écritures

IDENTIFICATION DES PROBLÈMES

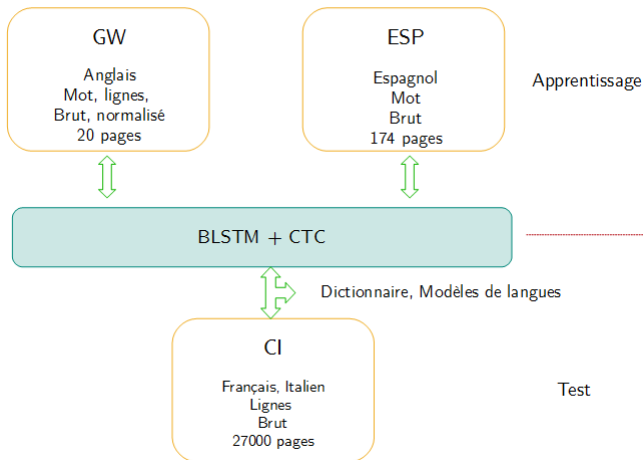


FIGURE – Utilisation des ensembles de données pour l'apprentissage et le test du système

Transcription à partir du BLSTM+CTC : sujet aux erreurs

Vocabulaire fermé :

- mot hors-vocabulaire

- variations de formes entre *Repertorio* et registres

Zone *titre* dans les registres :

- plusieurs lignes

- manque des mots

- titres de ballets et d'opéras manquants – ex : "isle sonnante" de Rabelais

CORPUS ET CARACTÉRISTIQUES

Au format xml ⇒ créer les images des lignes mots

174 pages de registres de mariage

5447 lignes

56378 imagettes de mots

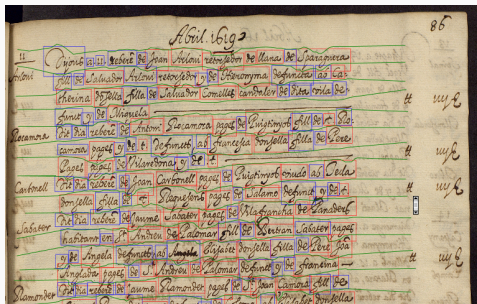
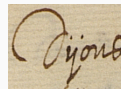


FIGURE – Exemple d'un document



(a) Dijous



(b)

Arlovi



(c)

de

FIGURE – Imagettes des mots

CORPUS GEORGES WASHINGTON (GW)

Couple de fichiers (.tiff .gtp) ⇒ créer les images des mots **sans normalisation**

20 pages de lettres

2919 imagerie de mots

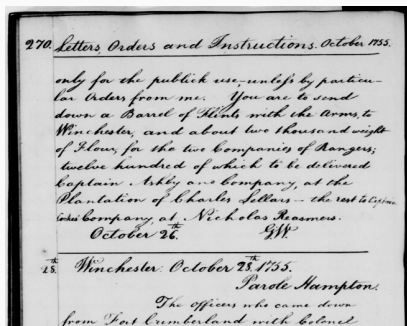


FIGURE – Exemple d'une page de GW (2700270)

(a) hundred

(b) are

(c) GW

FIGURE – Imagerie des mots du registre 2700270

Les caractéristiques statistiques (dans une colonne donnée) noté 4F :

- profils haut et bas : premier et dernier pixels d'encre

- nombre de pixels d'encre

- nombre de transitions entre l'encre et le fond

Les caractéristiques directionnelles : Histogramme des Gradients
Orientés noté 8HOG, 64HOG

- Sur GW normalisé

- cell : (120,4) et (15,4)

- block : 1,1

PREMIERS TESTS ET RÉSULTATS

Données : répartition aléatoire

train : 90% (4500)

validation : 5% (190)

test : 5% (200)

⇒ 5 samples : pour observer l'apprentissage des sorties du réseau

Configuration :

type de données : mots normalisés

caractéristiques : 4F, 8HOG, 64HOG

epoch : 40

lstm : 100 nodes

output : 75 caractères + blank

décodage : le meilleur chemin

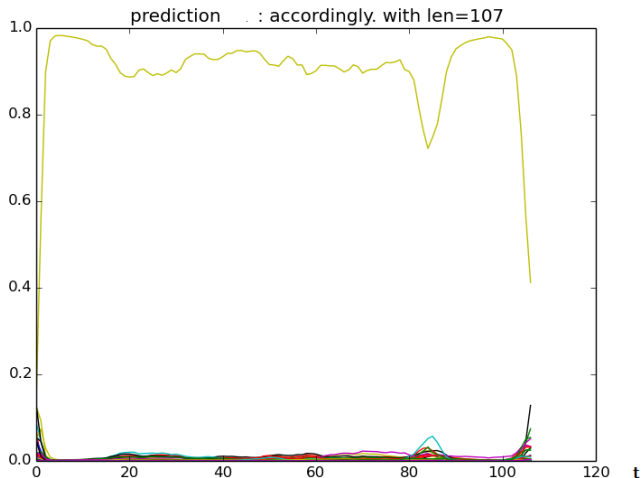


FIGURE – Courbe de prédiction du mot "accordingly." à 0 epoch

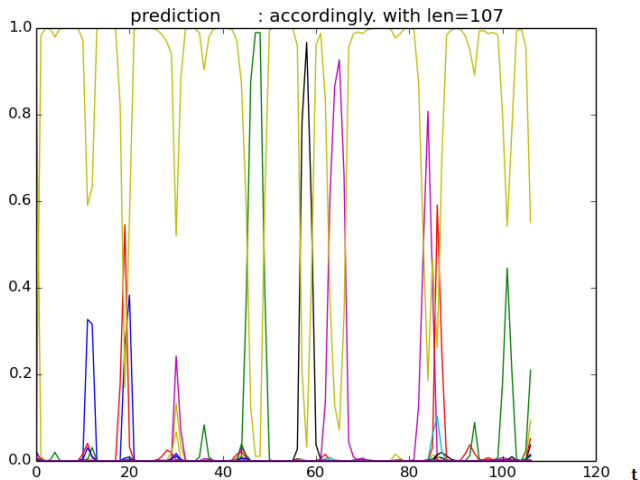


FIGURE – Courbe de prédiction du mot "accordingly." à 10 epoch

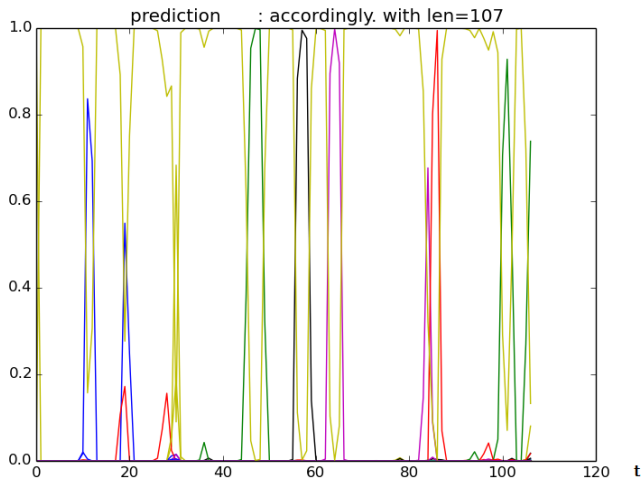


FIGURE – Courbe de prédiction du mot "accordingly." à 20 epoch

$$\text{WER} = \frac{\text{ins} + \text{supp} + \text{sub}}{\text{ins} + \text{supp} + \text{sub} + \text{corr}}$$

TABLE – Résultats obtenus et comparaison avec l'état de l'art

Méthode	Dataset	# feat	WER
BLSTM-CTC	IAM-OnDB	-	26%
BLSTM-CTC+ML	IAM-OnDB	-	20.4%
BLSTM	RIMES (fr,59203)	-	60.5%
BLSTM-CTCMMC	RIMES (fr,59203)	64 (pixels)	13.2%
BLSTM-CTCMMC	RIMES (fr,59203)	64 (HOG)	12.5%
BLSTM-CTC	GW (en,4893)	4 (4F)	14.7%
BLSTM-CTC	GW (en,4893)	8 (HOG)	47.39%
BLSTM-CTC	GW (en,4893)	64 (HOG)	13.5%
BLSTM-CTC	GW (en,4893)	64 (15 HOG)	33.4%

REFERENCES



Alex Graves, Santiago Fernández, Faustino Gomez, and Jürgen Schmidhuber.
Connectionist temporal classification : labelling unsegmented sequence data with recurrent neural networks.

In Proceedings of the 23rd international conference on Machine learning, pages 369–376. ACM, 2006.



Alex Graves.

Sequence transduction with recurrent neural networks.

arXiv preprint arXiv :1211.3711, 2012.



Andreas Fischer, Markus Wüthrich, Marcus Liwicki, Volkmar Frinken, Horst Bunke, Gabriel Viehhauser, and Michael Stolz.

Automatic transcription of handwritten medieval documents.

In Virtual Systems and Multimedia, 2009. VSMM'09. 15th International Conference on, pages 137–142. IEEE, 2009.



Théodore Bluche, Hermann Ney, Jérôme Louradour, and Christopher Kermorvant.

Framewise and ctc training of neural networks for handwriting recognition.

In Document Analysis and Recognition (ICDAR), 2015 13th International Conference on, pages 81–85. IEEE, 2015.



Nobuyuki Otsu.

A threshold selection method from gray-level histograms.

Automatica, 11(285-296) :23–27, 1975.



Kai Chen, Mathias Seuret, Marcus Liwicki, Jean Hennebert, and Rolf Ingold.

Page segmentation of historical document images with convolutional autoencoders.

In Document Analysis and Recognition (ICDAR), 2015 13th International Conference on, pages 1011–1015. IEEE, 2015.