



**HAL**  
open science

# Hybridation d’approches “ transparentes ” et basées “ Deep Learning ” pour l’analyse automatisée de productions graphiques d’élèves dans le contexte de l’éducation

Simon Corbillé, Eric Anquetil, Elisa Fromont

## ► To cite this version:

Simon Corbillé, Eric Anquetil, Elisa Fromont. Hybridation d’approches “ transparentes ” et basées “ Deep Learning ” pour l’analyse automatisée de productions graphiques d’élèves dans le contexte de l’éducation. SIFED 2020 - Symposium International Francophone sur l’Ecrit et le Document, Jul 2020, En ligne, France. hal-03992931

**HAL Id: hal-03992931**

**<https://hal.science/hal-03992931v1>**

Submitted on 16 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hybridation d'approches « transparentes » et basées « Deep Learning » pour l'analyse automatisée de productions graphiques d'élèves dans le contexte de l'éducation.

SIFED- 2020

---

Simon Corbillé

Directeurs de thèse : Éric Anquetil, Élisabeth Fromont  
Laboratoire IRISA – Équipes IntuiDoc, Lacodam

# Sommaire

---

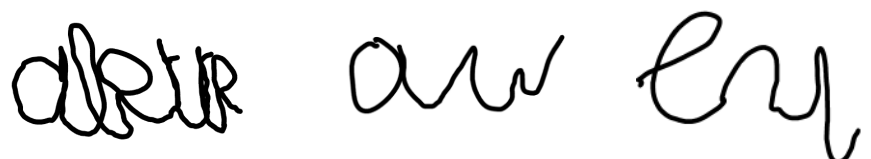
- Introduction
- Travaux réalisés
- Travaux en cours
- Perspectives

# Problématiques

---


- Ecriture manuscrite d'enfant
  - En cours d'apprentissage
  - Approximative
- Reconnaissance
- Analyse
  - Formation de la lettre
  - Fautes d'orthographe
  - Score par rapport à la consigne

alette                      an                      lu



Exemples écriture déformée

beau                      an                      ours                      vert

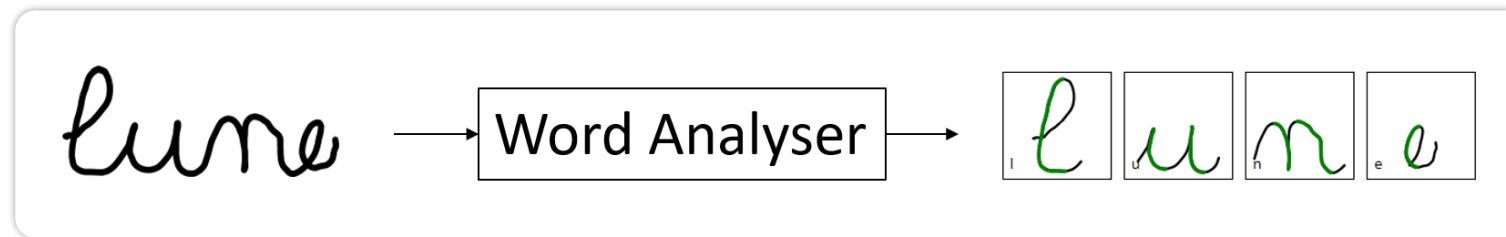


Exemples fautes d'orthographe

Sujet de thèse : Hybridation d'approches « transparentes » et basées « Deep Learning » pour l'analyse automatisée de productions graphiques d'élèves dans le contexte de l'éducation.

# Systeme existant : reconnaissance et analyse de mot [1]

- Laboratoire commun ANR : Script&Labs
- Entrées :
  - Tracé manuscrit
  - Consigne

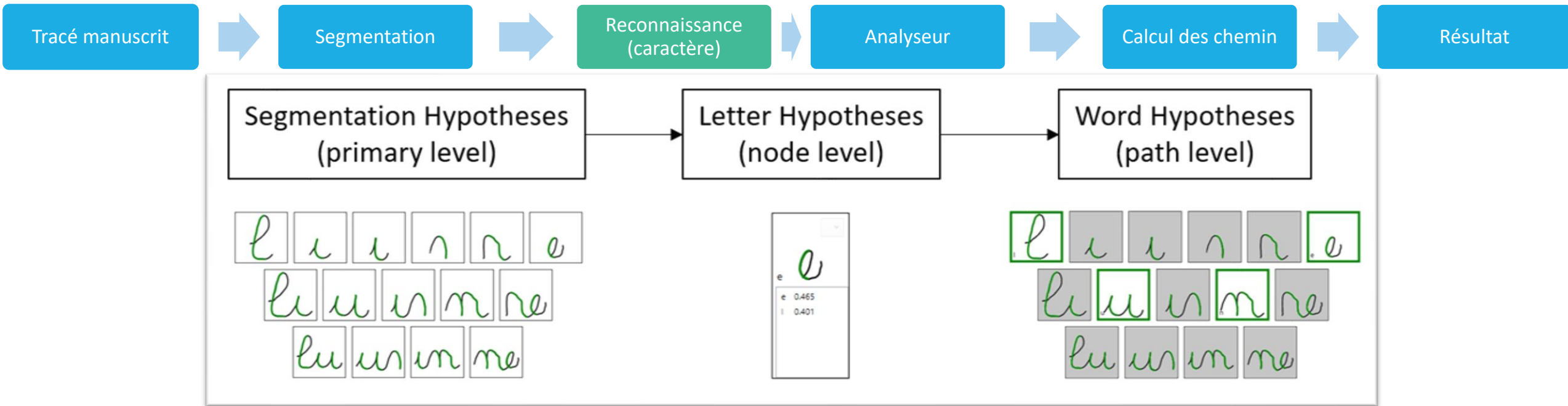


Exemple avec le mot lune

[1] Simonnet et al. Evaluation of Children Cursive Handwritten Words for e-Education. Pattern Recognition Letters 2019

# Processus de reconnaissance et d'analyse de mot

- Découpage en treillis
  - Connaissances liées à l'écriture

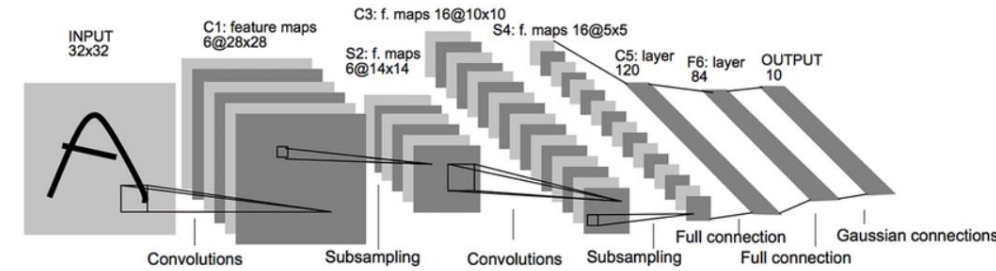
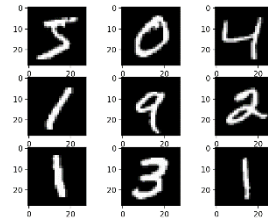


# Etudes comparatives architectures CNN

- But : connaître les performances obtenues avec différents réseaux

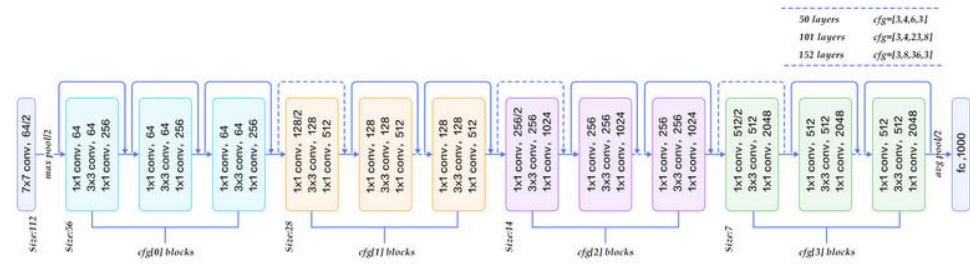
- LeNet [1]

- Reconnaissance de chiffres jeu de données MNIST
- ~60 000 paramètres



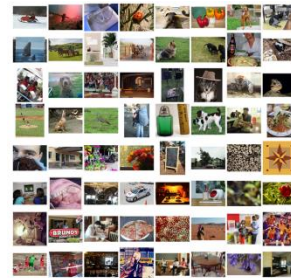
- ResNet [2]

- Connexions résiduelles
- ~11 000 000 paramètres



- VGG [3]

- Reconnaissance d'images
- ~28 000 000 paramètres



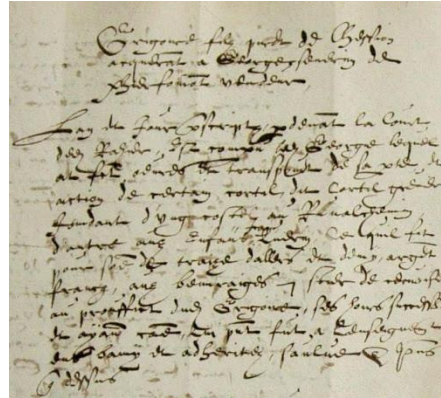
[1] Lecunet al. Gradient based learning applied to document recognition. Proceedings of the IEEE 1998.

[2] He et al. Deep residual learning for image recognition. CVPR 2016

[3] Simonyan et al. Very deep convolutional networks for large-scale image recognition. ICLR 2015.

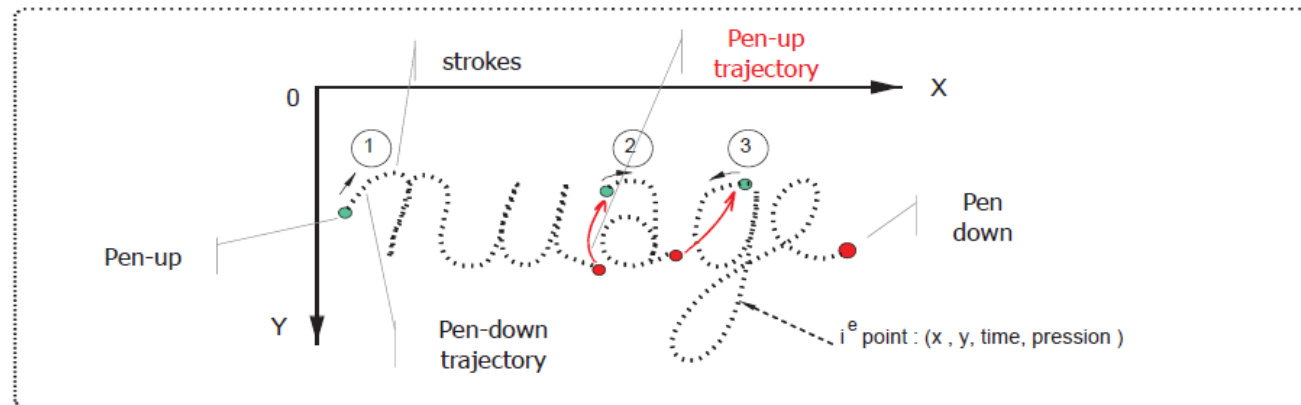
# Reconnaissance hors ligne et en ligne

- Hors ligne
  - Reconnaissance d'image



„Mr. Rowell finds it easier to take it out of mothers, childrens and sick people than to take on this vast industrie,” Mr. Bonn commented icily. „let us have a full inquiry into the cost of drugs and the pharmaceutical industrie.” The health of children today owed much to the welfare food scheme. It was maintained during the war. Now in conditions of Tory affluence it seemed it could not be carried on.

- En ligne
  - Série de points
    - X, y, pression, temps





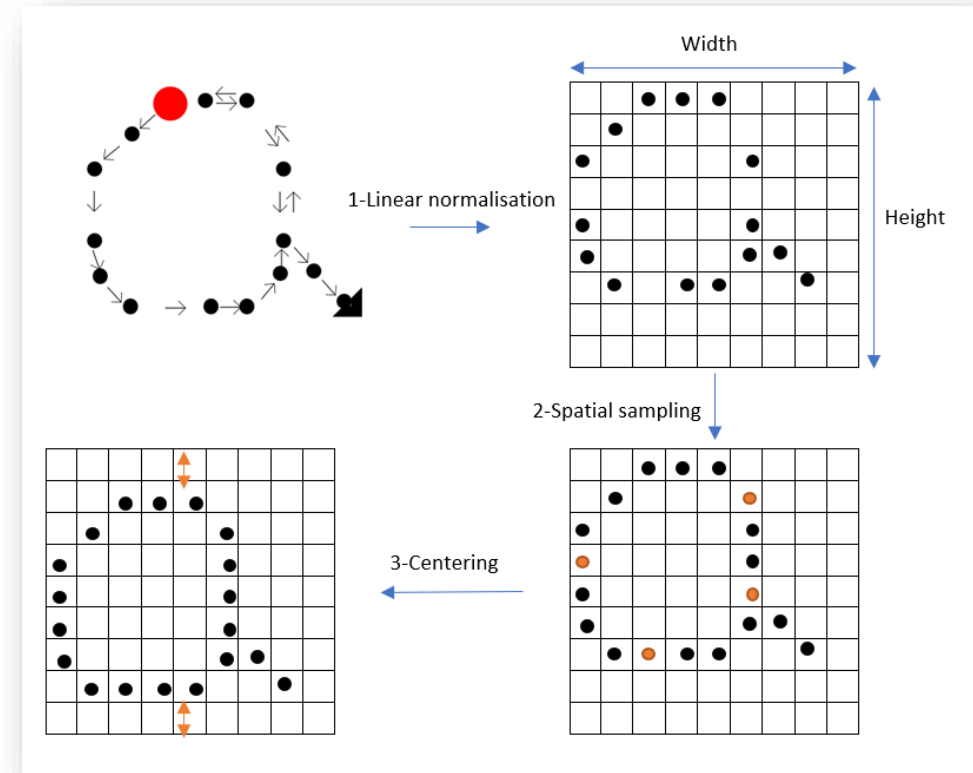
# Sommaire

---

- Introduction
- **Travaux réalisés**
- Travaux en cours
- Perspectives

# Encodage des données

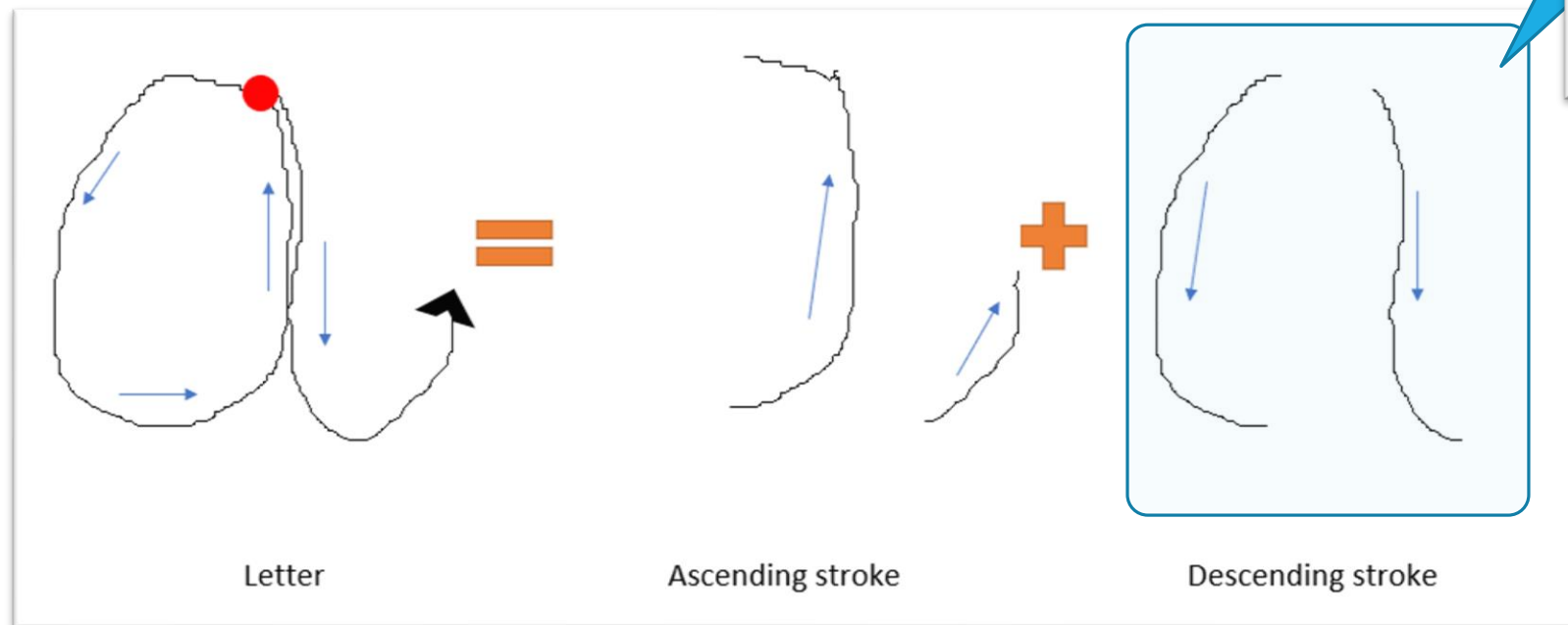
- Prétraitement
- Conversion en matrice de valeurs
  - 1 s'il y a un tracé
  - 0 sinon



Conversion du signal en ligne vers une image

# Information dynamique : la direction

- Extraction des zones descendantes
- Caractérise la direction du tracé

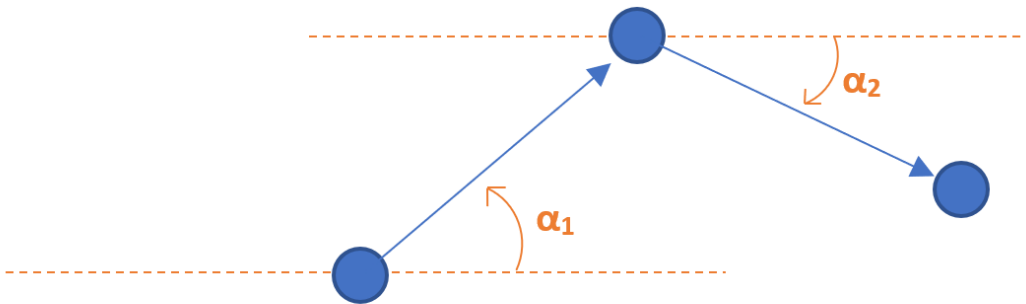


		1					
	1			1			
1				1			
1				1			
1				1			
1				1	1		
	1						1

Décomposition d'un tracé en zones ascendantes et descendantes

# Information sur l'orientation

- Utilisation de l'angle entre un point et le point suivant
  - Cosinus
  - Sinus



Calcul de l'angle

		-0,8	-1	-1		
	-0,8			0,8		
0,2				0,2		
0,2				0,2		
0,2				0,2		
-0,8				0,2	-0,8	
1	1	1	0,8			-0,8

		-0,8	0,2	0,2		
	-0,8			-0,8		
-1				-1		
-1				-1		
-1				-1		
0,8				-1	-0,8	
0,2	0,2	0,2	0,8			-0,8

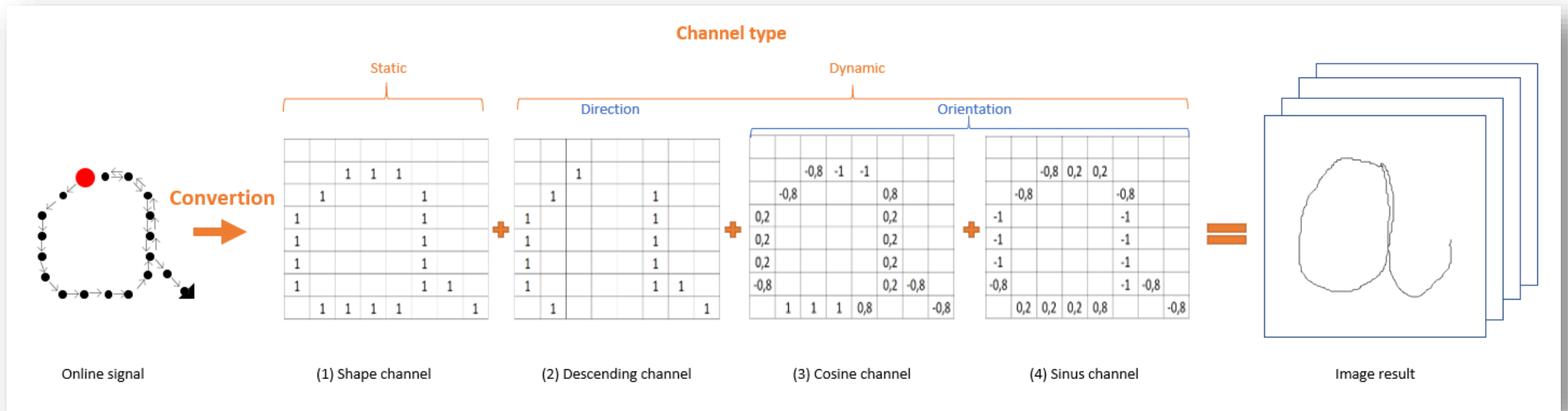
Stockage du cosinus et du sinus

# Deux approches de fusion

---

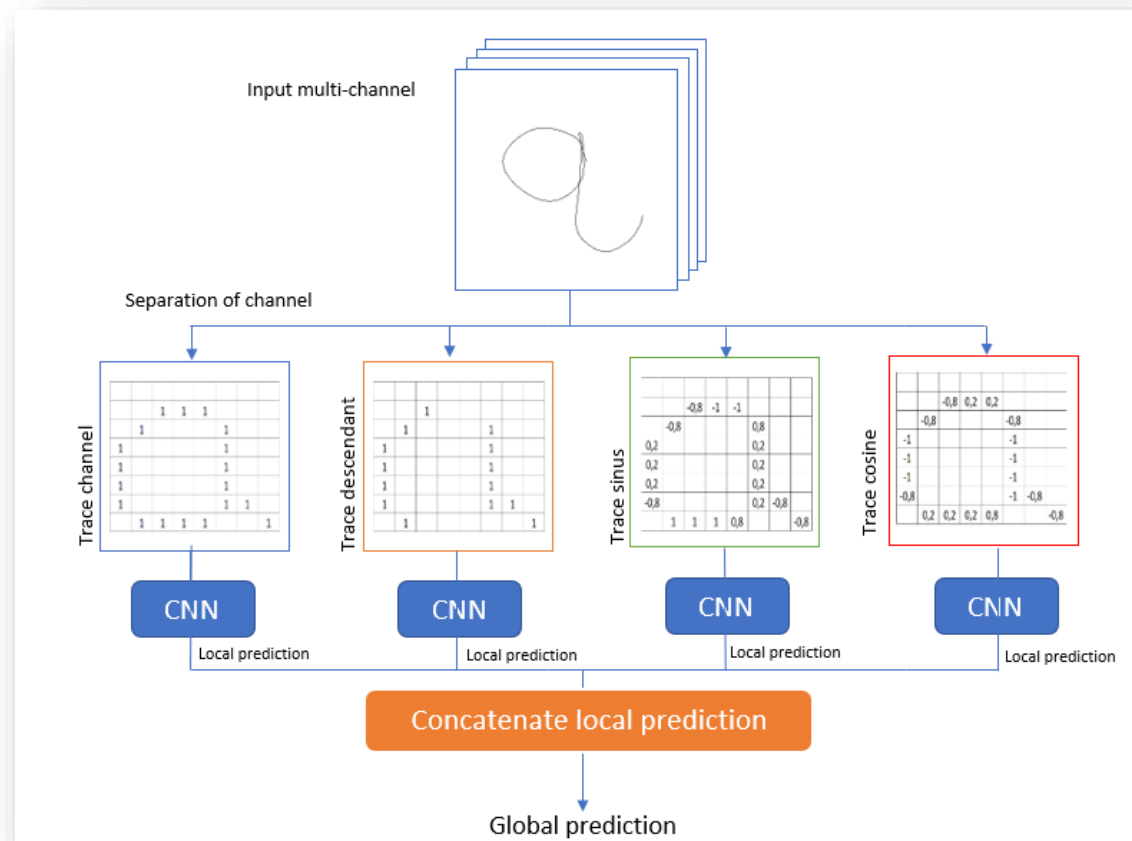
- Fusion des informations
  - Forme de la lettre
  - Direction
  - Orientation
  
- Fusion précoce
  - 1 image multicanal -> un réseau
  
- Fusion tardive
  - 4 canaux -> 4 images -> 4 réseaux

# Approche multicanal



Processus de conversion d'un tracé en ligne vers une image multicanal

# Approche ensemble de classifieurs



Architecture ensemble de classifieurs

# Les données

- Ecriture en ligne d'enfants

- Synthèse de données

- Déformation image
- Déformation en ligne

- Jeu de données

- Avant synthèse

- Apprentissage : 14 753
- Validation : 2600
- Test : 7086

- Après synthèse

- Apprentissage : 140 000
- Validation : 2 600
- Test : 7 086





# Expérimentations

---

- Objectifs
  - Evaluer les performances des différents réseaux
  - Impact de l'ajout d'information dynamique
  - Comparer la fusion précoce et tardive

# Résultats fusion précoce

- Type d'entrée

- A
  - Statique : forme de la lettre
- B
  - Statique : forme de la lettre
  - Dynamique : orientation (cosinus et sinus)
- C
  - Statique : forme de la lettre
  - Dynamique : direction (zone descendante)
- D
  - Statique : forme de la lettre
  - Dynamique
    - Orientation (cosinus et sinus)
    - Direction (zone descendante)

Input Config	Static	Static + Dynamic		
	A	B	C	D
<b>Lenet-5</b>				
Accuracy	90.62	91.35	91.93	<b>92.66</b>
Variance	0.0240	0.1599	0.0121	0.0049
<b>ResNet-18</b>				
Accuracy	94.05	94.06	94.33	<b>94.64</b>
Variance	0.0100	0.1764	0.0030	0.0030
<b>VGG-11</b>				
Accuracy	94.54	94.64	94.84	<b>95.00</b>
Variance	0.0066	0.0580	0.0323	0.0042

## ICFHR 2020 :

- Simon Corbillé, Élisabeth Fromont, Éric Anquetil, Pauline Nerdeux. Fusion of multiple channels with convolutional neural networks for online children handwriting recognition. **ICFHR 2020**

# Résultats fusion tardive

- Type d'entrée
  - 1 : statique - forme de la lettre
  - 2 : dynamique - zones descendantes
  - 3 : dynamique - cosinus
  - 4 : dynamique - sinus

Input Config	Trace	Dynamic			Ensemble
	Type 1	Type 2	Type 3	Type 4	
LeNet-5					
Accuracy	91.10	77.49	90.78	88.15	<b>93.35</b>
ResNet-18					
Accuracy	93.35	74.44	93.60	93.19	<b>94.87</b>
VGG-11					
Accuracy	94.24	75.30	94.24	93.77	<b>94.76</b>

# Sommaire

---

- Introduction
- Travaux réalisés
- Travaux en cours
- Perspectives

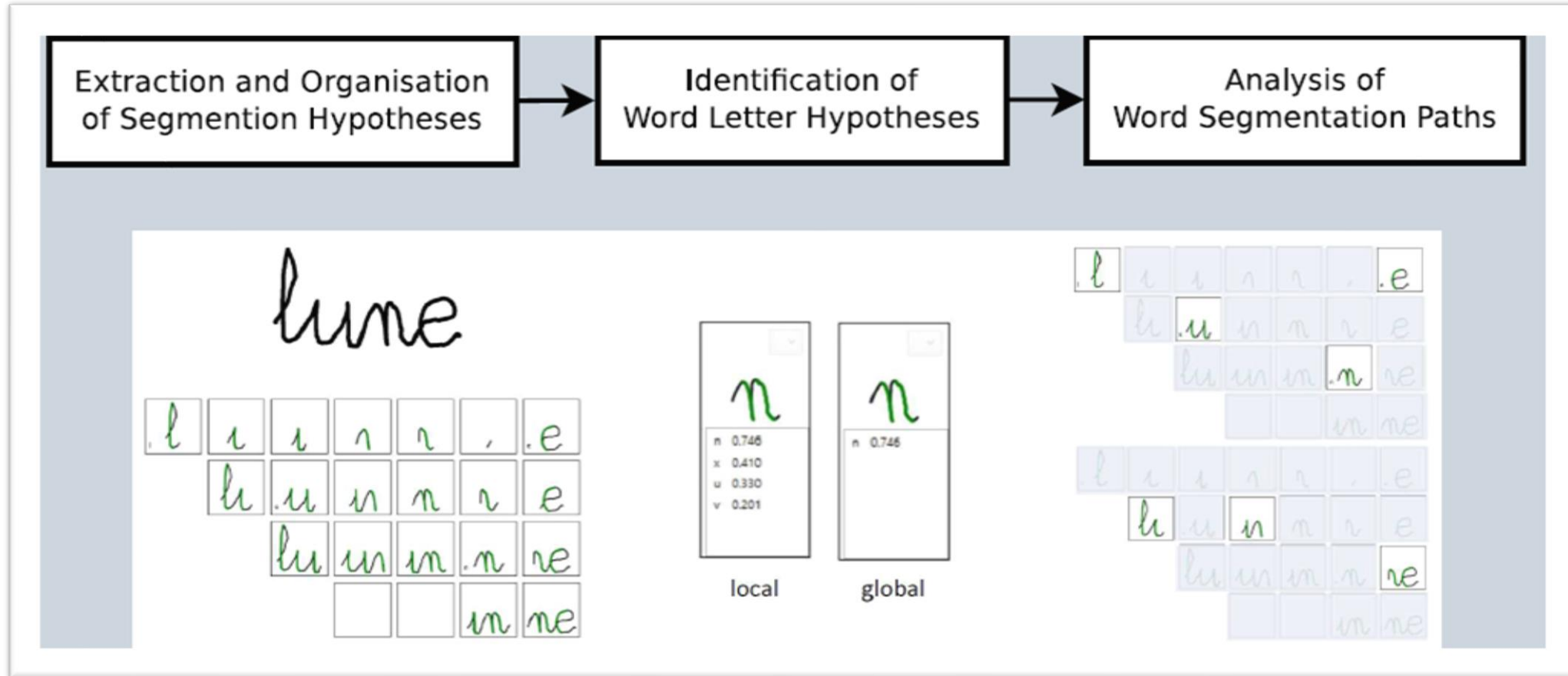
# Intégration reconnaissseur de caractère

---

- Précision
  - Actuel : 93.75 %
  - Nouveau : 95.00 %
- Objectif : impact sur le reconnaissseur de mot

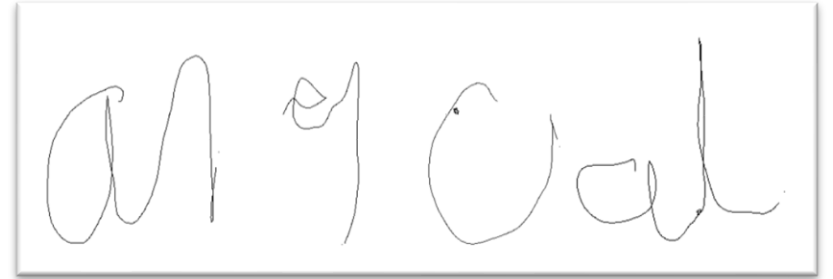


# Systeme existant



# Notion de rejet

- Ajout d'une classe rejet au 26 lettres de l'alphabet
  - Rejet = lettre + morceau d'autre lettre
- Classe rejet
  - Apprentissage : 13662
  - Validation : 2000
  - Test : 10608



Exemple rejet

- Résultats avec une architecture VGG :
  - Accuracy : 81.64% (variance 0.4397)

Vrai positif <b>99.70%</b>	Faux positif <b>26.97%</b>
Faux négatif <b>0.30%</b>	Vrai négatif <b>73.03%</b>

# Sommaire

---

- Introduction
- Travaux réalisés
- Travaux en cours
- Perspectives



# Perspectives

---

- Analyse du tracé
  - Expliquer pourquoi le réseau à pris cette décision
  - Déformation, faute d'orthographe
- Réseau multi-dimensionnels
  - Segmentation
  - Reconnaissance
- Système end to end
  - Reconnaissance et analyse de mots, phrase
- Approche de fusion
- Série temporelle multivariée