

Projet de deep learning

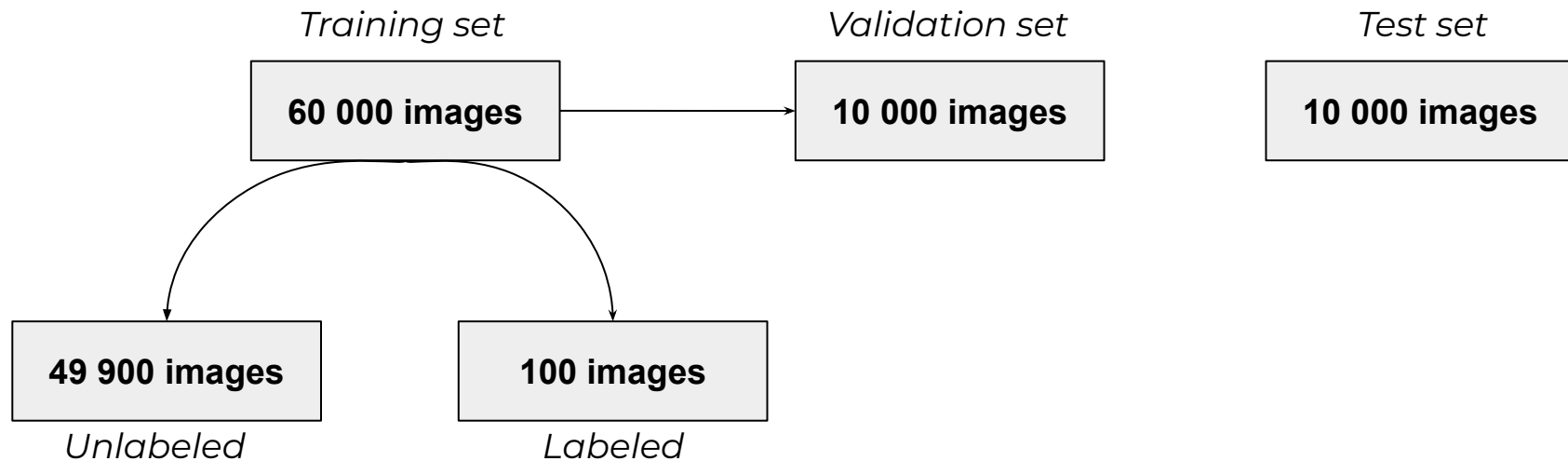
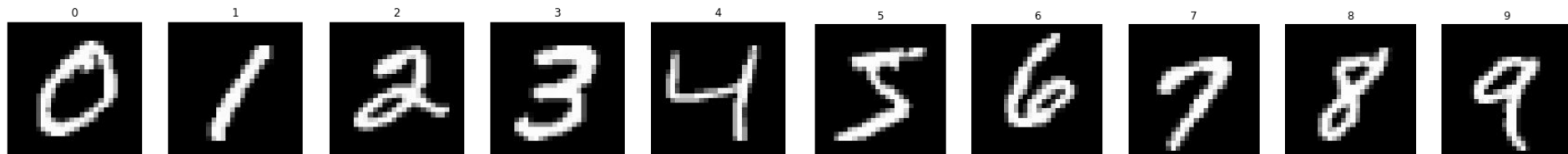
Noisy Student

Etienne GAUCHER

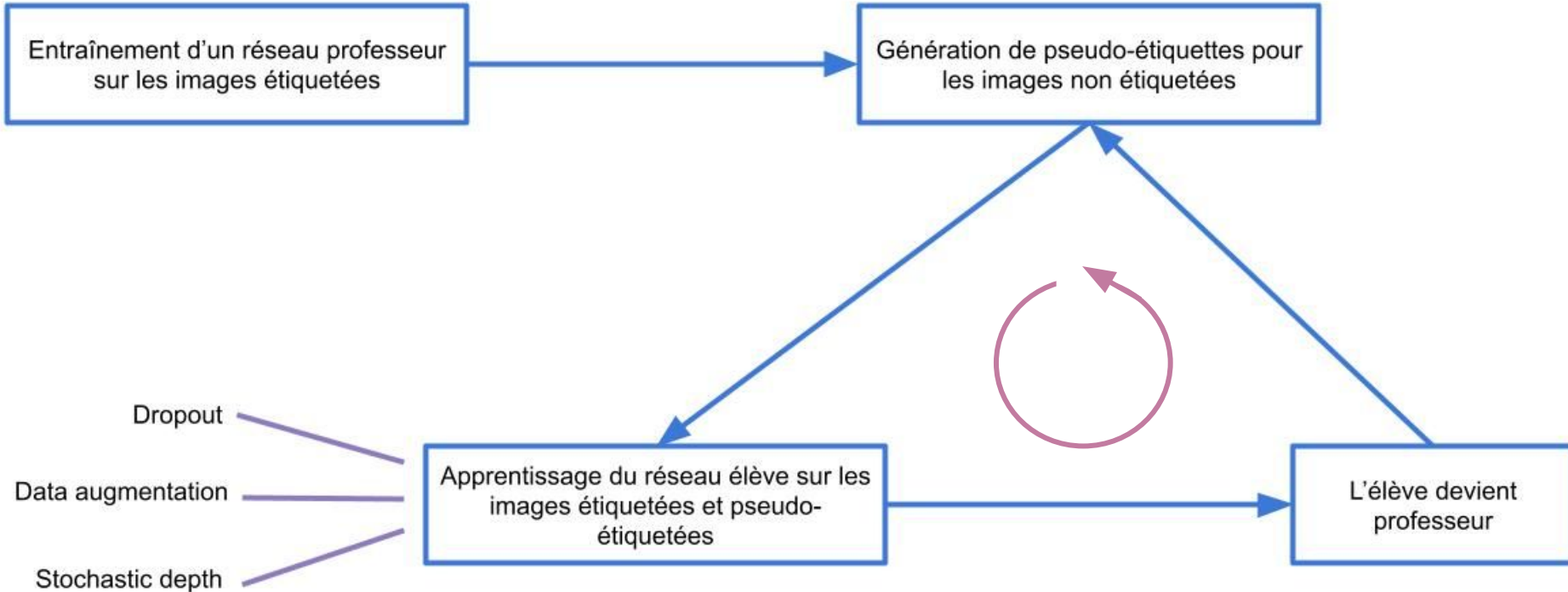
Sommaire

- | | |
|------------------------------|------------------------------------------------------|
| I. Introduction | VI. Modification des labels |
| II. La méthode Noisy Student | VII. Résultats des modèles : Dropout,
RandAugment |
| III. Les types de “bruit” | VIII. Conclusion |
| IV. Modèle de base | |
| V. Early Stopping | |
-

I. Introduction

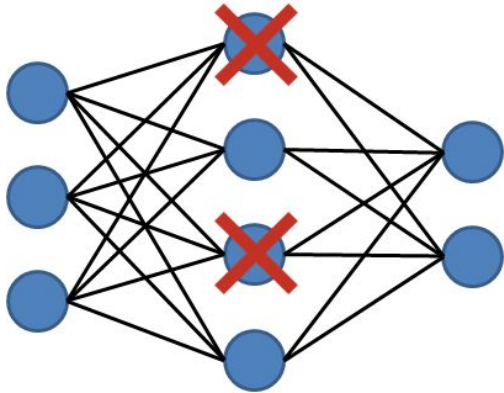


II. La méthode Noisy Student

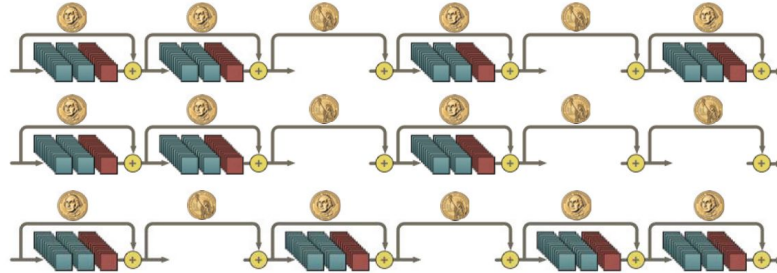


III. Les types de “bruit”

DropOut



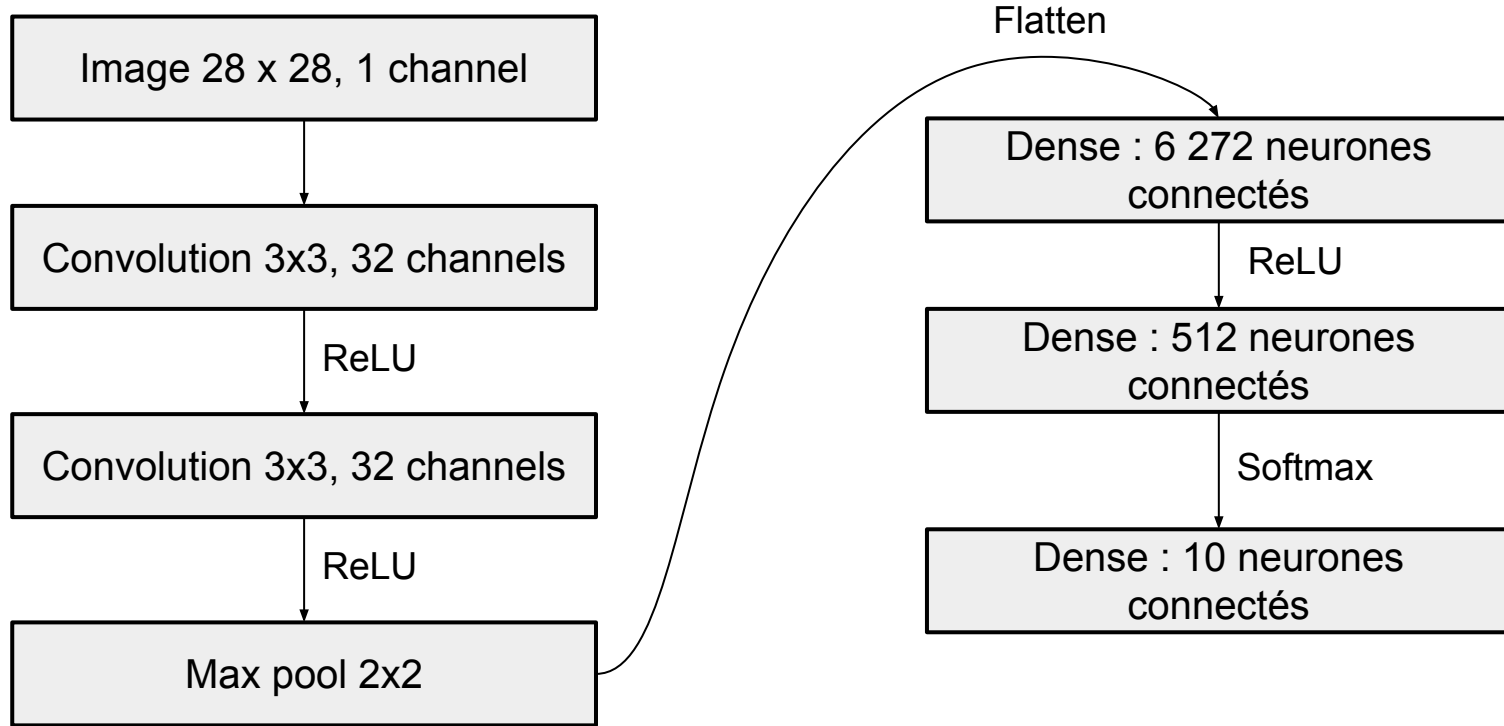
Stochastic Depth



Random Augment



IV. Modèle de base

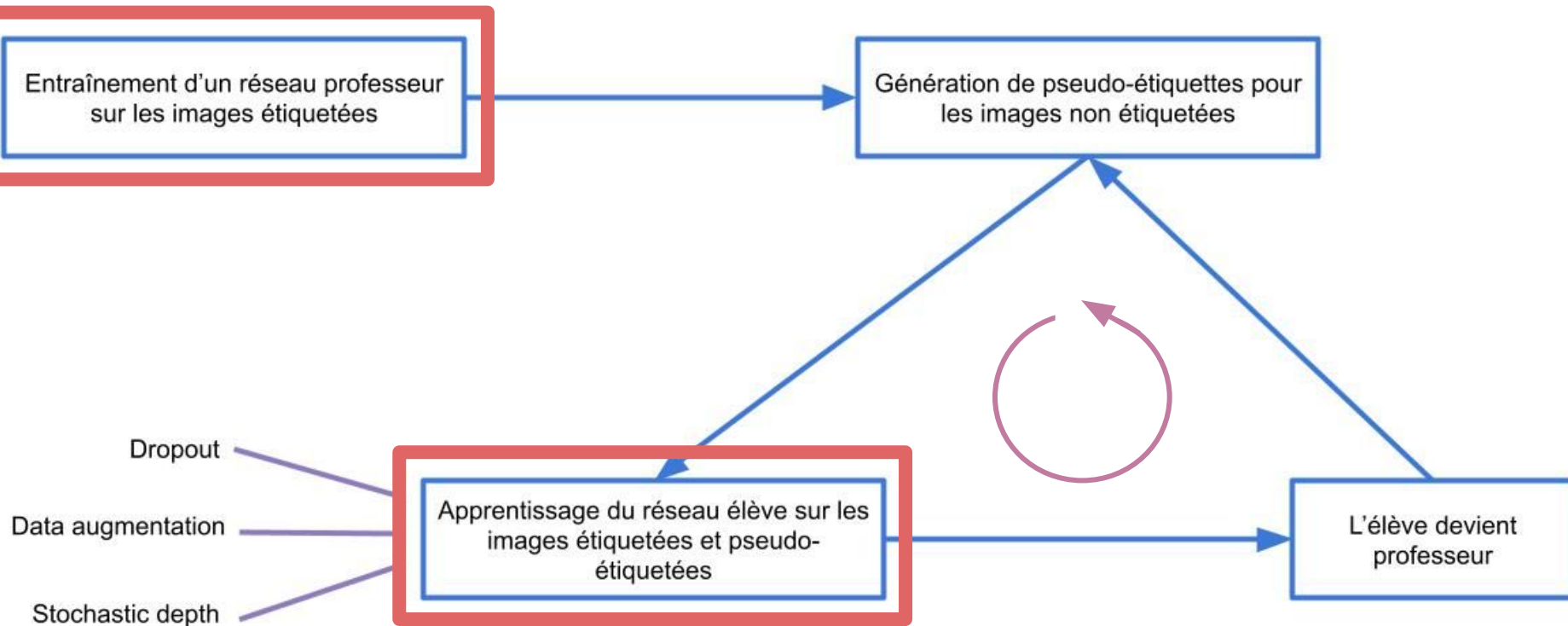


V. Early Stopping

2 types d'Early stopping :

- Sur l'apprentissage des réseaux professeur et élève

V. Early Stopping

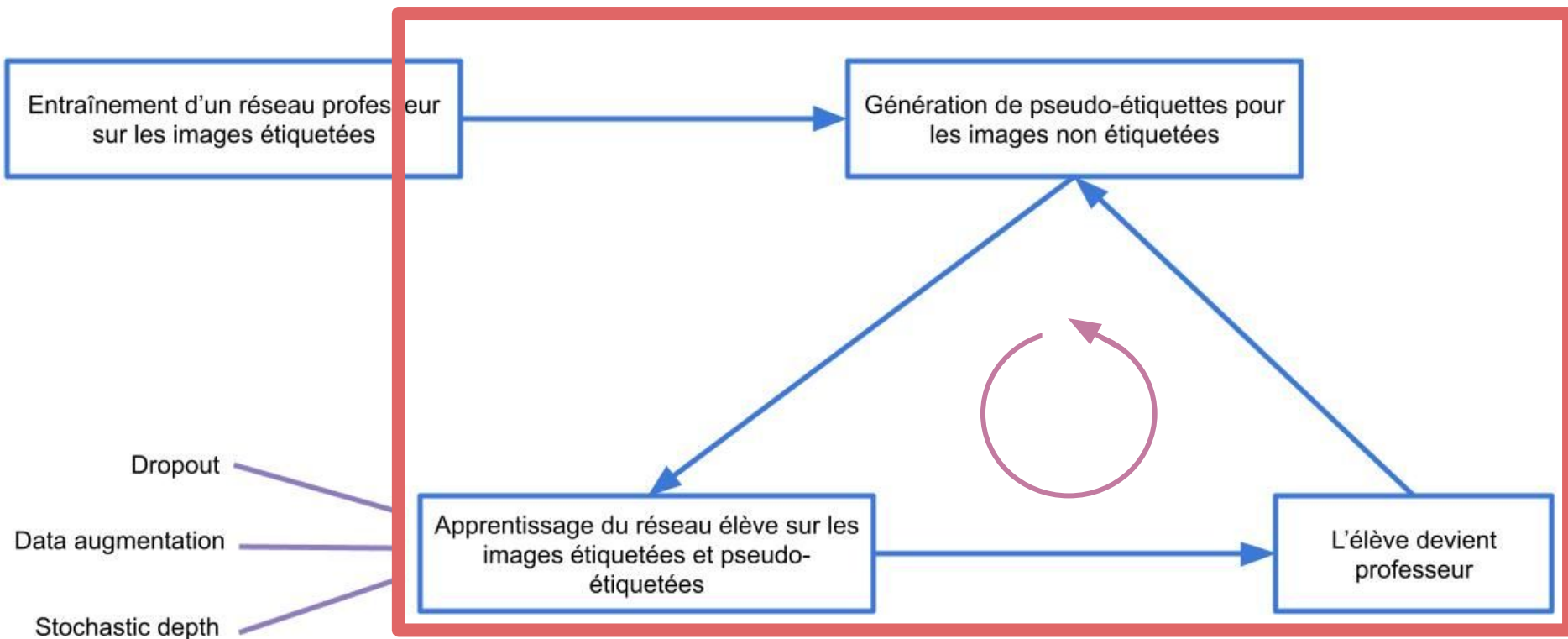


V. Early Stopping

2 types d'Early stopping :

- Sur l'apprentissage des réseaux professeur et élève
 - Sur le nombre de boucles professeur-élève
-

V. Early Stopping



VI. Modification des labels

Modification des labels dans le dataset



Modification des labels dans le dataloader

VII. Résultats des modèles

VII.A Elève sans bruit

Numéro de simulation	1	2	3	4	5
Baseline (accuracy)	0.720	0.810	0.760	0.795	0.766
Meilleure amélioration	0.725	0.817	0.774	0.797	0.756
Différence	0.5 %	0.7 %	1.4 %	0.2 %	-1.0 %
Nombre boucle Professeur-Elève	1	1	4	1	1

VII. Résultats des modèles

VII.B Elève avec DropOut

Numéro de simulation	1	2	3	4	5
Baseline (accuracy)	0.772	0.779	0.738	0.822	0.785
Meilleure amélioration	0.806	0.766	0.740	0.84	0.80
Différence	3.4 %	-1.3 %	0.2 %	1.8 %	1.5 %
Nombre boucle Professeur-Elève	3	1	1	2	1

VII. Résultats des modèles

VII.C Elève avec Dropout et RandAugment

Numéro de simulation	1	2	3
Baseline (accuracy)	0.782	0.753	0.791
Meilleure amélioration	0.884	0.882	0.930
Différence	10.2 %	12.9 %	13.9 %
Nombre boucle Professeur-Elève	5	8	4

VIII. Limites

- On effectue peu de simulations
 - D'autres méthodes/modèles sont à tester :
 - Filtration et stratification des images pseudo-étiquetées
 - Pondération des fonctions de perte
 - Modèle de base avec des couches additionnelles
 - Modèle avec réseaux résiduels pour voir les effets du Stochastic Depth
 - Transfer-learning du modèle de l'article
-