



CentraleSupélec

Rapport sur la thèse présentée par Quentin Ferdinand
Mitigating catastrophic forgetting via feature transfer and knowledge consolidation for deep-
incremental learning

En vue de l'obtention du titre de

Docteur de l'École Nationale Supérieure de Techniques Avancées Bretagne

École doctorale n° 648 : Sciences pour l'Ingénieur et le Numérique

Spécialité : Sciences et Technologie de l'information et de la communication - Informatique

Préparée au Lab-STICC(UMR CNRS 6285)

par Céline Hudelot, Professeur des Universités

Laboratoire MICS - CentraleSupélec- Université Paris Saclay

Contexte

La thèse de Quentin Ferdinand se place dans le cadre d'une collaboration de type CIFRE entre la société Naval Group, spécialisée dans la construction navale de défense et le laboratoire Lab-STICC. Elle adresse la problématique de l'apprentissage automatique incrémental, c'est-à-dire la conception de modèles d'apprentissage capables de s'adapter à de nouvelles données et de mettre à jour leurs connaissances de manière continue et incrémentale, sans avoir besoin de reprendre l'apprentissage depuis le début. C'est un sujet de recherche très actuel et très compétitif car il répond à de nombreux contextes applicatifs. Dans sa thèse, Quentin Ferdinand s'intéresse principalement à la classification incrémentale d'images.

Dans ce contexte, Quentin Ferdinand propose plusieurs contributions expérimentales et méthodologiques pour minimiser le problème bien connu de l'oubli (catastrophic forgetting) des connaissances sur les anciennes classes afin de maintenir un équilibre entre l'apprentissage des nouvelles classes et la consolidation des connaissances antérieures. Il propose notamment une contribution qui permet d'exploiter l'apprentissage contrastif, très utilisé en apprentissage de représentations, dans le contexte de la classification incrémentale. Il propose aussi une contribution originale basé sur un modèle d'apprentissage neuronal dynamique et une stratégie d'expansion et de compression de modèles.

Étude du mémoire

Quentin Ferdinand propose un manuscrit de 116 pages, rédigé en anglais, articulé en 4 chapitres intercalés entre une introduction et une conclusion générale. Le manuscrit propose une liste bibliographique de 79 références. Le manuscrit est clair et de bonne facture.

Le manuscrit commence par une introduction générale qui présente très brièvement le contexte scientifique et applicatif des travaux de thèse et qui introduit les différentes contributions. Il aurait pu être intéressant ici de détailler et caractériser un peu plus le contexte applicatif et ses défis scientifiques qui vont, il me semble, au-delà de la classification incrémentale d'images naturelles. Cependant, l'introduction permet de bien comprendre le positionnement des travaux de Quentin Ferdinand.

Le **chapitre 2** présente un état de l'art relatif à la classification d'images par apprentissage et apprentissage profond et à l'apprentissage incrémental. Après un rapide rappel des principes de l'apprentissage supervisé pour la classification d'images et des réseaux convolutionnels, Quentin Ferdinand présente un état de l'art relativement complet de l'apprentissage incrémental en commençant par présenter ses défis scientifiques (l'oubli, l'occupation en mémoire, le temps d'apprentissage et le compromis plasticité/stabilité), les différents scénarios de sa mise en œuvre et une formulation mathématique du problème. Une suggestion ici est qu'il aurait pu être intéressant de décomposer ici le modèle en un extracteur de caractéristiques et un classifieur comme cela est souvent fait en classification d'images et pour permettre de mieux mettre en avant la représentation. Il présente ensuite, de manière

détaillée, les principales familles d'approches en apprentissage incrémental : celles basées sur la sauvegarde ou la répétition du passé (rehearsal based) et celles basées sur la régularisation et la distillation de connaissance. Il termine avec la prise en compte du déséquilibre et les approches permettant de corriger le biais vers les nouvelles classes et il conclut le chapitre en présentant l'approche qui servira de référence dans la suite de ses travaux.

Dans le **chapitre 2**, Quentin Ferdinand présente un ensemble de contributions expérimentales menées au début des travaux de thèse et visant à mieux comprendre les défis liés à l'oubli en apprentissage incrémental et les limites des différentes approches de la littérature. Il s'intéresse notamment à l'espace des caractéristiques dans lequel il étudie la représentativité du centre de gravité de chaque classe et son voisinage et les distances entre classes et intra classes. Les limites observées sont souvent dues aux hypothèses sous-jacentes portant sur l'espace de représentation des approches de l'état de l'art. Il aurait pu être intéressant de mieux expliciter ces différentes hypothèses. De même, il aurait pu être intéressant de faire un parallèle avec des travaux connexes, comme en apprentissage few shot pour lequel les notions de prototypes (instance moyenne) ont été très largement étudiées.

Quentin Ferdinand propose ensuite de s'intéresser aux différentes stratégies de sélection des exemples à garder en mémoire ainsi qu'aux différents types de distillation de connaissances. Les expériences sont menées sur le jeu de données CIFAR 100 avec un modèle ResNet-32. Sur ce jeu de données, les résultats montrent l'importance de l'espace des caractéristiques et les avantages et les limitations des différentes fonction de perte de distillation de connaissances. Une observation intéressante concerne aussi ici le faible impact de la stratégie de sélection des exemples à mettre en mémoire sur les performances. Dans la mesure où ces expériences n'ont été menées que sur un seul jeu de données et avec un seul modèle backbone, il est difficile de tirer des conclusions générales sur les observations faites. Ces dernières mettent cependant en avant, en phase avec la recherche actuelle en reconnaissance visuelle, l'importance de la représentation et motivent la contribution suivante.

Le **chapitre 3** décrit une des premières contributions méthodologiques de la thèse. L'idée principale est de tirer parti de l'apprentissage contrastif très développé et très performant en apprentissage de représentation par auto-supervision. Après un bref rappel des bases de l'apprentissage contrastif non supervisé et supervisé et des travaux proches de l'état de l'art, Quentin Ferdinand introduit sa stratégie d'apprentissage. L'idée est de permettre, par l'apprentissage contrastif, d'apprendre une meilleure représentation (au sens améliorer sa transférabilité) à chaque étape incrémentale. Il introduit donc dans sa fonction objectif deux termes contrastifs. Le premier terme correspond à une adaptation de la fonction de perte proposée en apprentissage contrastif supervisé au contexte de l'apprentissage incrémental. Cette adaptation consiste à ne considérer, dans un mini-batch augmenté, que les instances des nouvelles classes comme exemples positifs et les exemples correspondant aux anciennes classes comme négatifs. D'autres stratégies auraient-elles pu être conçues et testées ? Ce terme contrastif favorisant les nouvelles classes, Quentin Ferdinand propose, pour compenser, d'ajouter aussi un terme contrastif supervisé pour la distillation de connaissances qui permet de conserver les similarités apprises pour les classes des états incrémentaux précédents. Quentin Ferdinand propose une validation expérimentale sur les jeux de données CIFAR-100 et ImageNet-100 avec un scénario à 10 étapes incrémentales de 10 classes qui sont tirées aléatoirement. La comparaison avec d'autres approches de l'état de l'art montre l'intérêt de l'approche proposée. Quentin Ferdinand propose aussi une étude d'ablation pour mieux comprendre les effets et les apports de chaque terme. Ici, il aurait aussi pu être intéressant de voir, dans la lignée de ce que Quentin Ferdinand propose dans le chapitre 2, si l'approche permet effectivement de mieux gérer l'oubli sous l'angle de l'étude des représentations. De même, une étude sur la nature des augmentations à utiliser pour améliorer la création de paires positives auraient pu être intéressante ici, notamment en essayant de plus prendre en compte le contexte applicatif de la thèse (type de classes, nature des données...).

Le **chapitre 4** décrit une deuxième contribution méthodologique de la thèse. Elle vise à répondre aux limitations computationnelles de l'approche contrastive en exploitant des architectures dynamiques et notamment le concept d'extracteur de caractéristiques extensible (approche DER) qui a montré une bonne capacité à gérer le compromis plasticité /stabilité. Quentin Ferdinand commence par rappeler le principe appliqué à l'apprentissage incrémental avec deux étapes : une étape d'expansion pour prendre

en compte les nouvelles tâches et une étape d'élagage pour limiter le temps de calcul et permettre le passage à l'échelle. Il s'intéresse ici aussi aux stratégies d'augmentation par mixage qui sont des stratégies bien connues pour permettre d'améliorer la généralisation des modèles et qui, pour certaines architectures permettent d'améliorer le non-oubli. Il propose notamment de l'appliquer entre deux images, une de la mémoire et une de la tâche incrémentale en cours. Il introduit ensuite l'approche FECIL qui applique le principe d'expansion et d'élagage sur l'extracteur de caractéristiques. En particulier, à chaque étape de l'apprentissage incrémental, on étend et on concatène l'extracteur précédent (gelé) avec un nouveau pour les nouvelles classes et on considère un classifieur étendu au nombre de classes total (nouvelles et anciennes). La fonction objectif est double avec un terme de classification (CE loss) et un terme d'apprentissage de représentation qui force l'apprentissage de caractéristiques discriminantes sur les nouvelles classes. Il propose ensuite, de manière originale de compresser le gros modèle obtenu par distillation plutôt que par masquage comme dans l'approche originale DER et il exploite une stratégie d'augmentation par mixage entre les classes passées et les nouvelles classes. Quentin Ferdinand propose une validation expérimentale sur les jeux de données CIFAR-100 et ImageNet-1000 sur différents scénarios incrémentaux et fournit une étude d'ablation portant sur les différentes composantes de la méthode. Les résultats montrent l'intérêt d'une part de l'approche FECIL mais aussi de la stratégie de mixage adaptée à l'apprentissage incrémental. Cette dernière contribution est une contribution très intéressante qui ouvre de nombreuses perspectives en apprentissage incrémental, notamment concernant les architectures dynamiques.

Le manuscrit se termine par une conclusion générale des travaux de thèse et une discussion sur quelques perspectives en apprentissage incrémental, notamment avec des architectures de type transformers.

Synthèse et conclusions de l'étude du mémoire de thèse en vue d'une soutenance

En résumé, Quentin Ferdinand présente, dans son manuscrit, un travail de recherche original en apprentissage incrémental avec deux contributions méthodologiques très intéressantes. Pour chaque contribution, Quentin Ferdinand adopte une méthodologie claire, argumentée et il l'évalue et la valide expérimentalement. Malgré les quelques commentaires et questions mis en avant dans ce rapport, la thèse de Quentin Ferdinand propose donc des avancées intéressantes dans le domaine de la classification incrémentale. Par ailleurs, les contributions ont donné lieu à plusieurs articles publiés dans des conférences internationales de bonne facture.

Ce manuscrit montre que Quentin Ferdinand a su acquérir pendant sa thèse toutes les bonnes pratiques nécessaires à l'exercice de la recherche :

- un examen et une analyse expérimentale de l'état de l'art, avec des références récentes, sur lesquelles il a su s'appuyer pour fonder ses propositions,
- la proposition d'hypothèses de travail bien fondées et de méthodes originales,
- la mise en œuvre d'expérimentations pour évaluer ses méthodes,
- et la capacité à clairement rendre compte de toutes ses contributions dans un document logiquement structuré.

Je donne donc un avis favorable à la soutenance de Quentin Ferdinand, en vue de l'obtention du Doctorat de l'École Nationale Supérieure de Techniques Avancées Bretagne.

Gif sur Yvette, le 08/11/2023

Céline Hudelot, Professeur en Informatique

CentraleSupélec, Laboratoire MICS, Mathématiques et Informatique pour la Complexité des Systèmes, 9 bis, 9 ter rue Joliot Curie - 91 190 Gif-sur-Yvette

e-mail : celine.hudelot@centralesupelec.fr



Céline Hudelot *C. Hudelot*