論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Balancing Old and New Classes for Class-Incremental Learning Methods
著者(和文)	JODELETQUENTIN MAGIC
Author(English)	Quentin Magic Jodelet
出典(和文)	学位:博士(学術), 学位授与機関:東京科学大学, 報告番号:甲第32号, 授与年月日:2024年12月31日, 学位の種別:課程博士, 審査員:村田 剛志,岡崎 直観,篠田 浩一,横田 理央,金崎 朝子
Citation(English)	Degree:Doctor (Academic), Conferring organization: Institute of Science Tokyo, Report number:甲第32号, Conferred date:2024/12/31, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

（博士課程）
Doctoral Program

論 文 要 旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	Computer Science Artificial Intelligence 系 コース	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of (Philosophy)
学生氏名： Student's Name	Jodelet Quentin Magic	審査員主査： Chief Examiner	Tsuyoshi Murata

要旨（英文 800 語程度）
Thesis Summary（approx.800 English Words ）

During the past decade, deep neural networks have led to numerous breakthroughs and reached state-of-the-art performances for various tasks of computer vision, natural language processing, or graph processing. However, while humans are able to continuously learn new knowledge throughout their lifetime, deep neural networks have to be trained on the complete dataset at once. In the Class-Incremental Learning setting, machine learning models are trained sequentially on new classes whereas training data from previously learned classes are no longer available. This training paradigm, closer to the human learning process, is challenging for deep neural network models due to catastrophic forgetting: when learning new classes, the model has a tendency to forget the old classes resulting in a significant deterioration of its performance.
This thesis focuses on Class-Incremental Learning for deep neural networks trained for image classification. In order to mitigate catastrophic forgetting, the author explores the important trade-off between old and new classes and proposes three distinct approaches meant to be seamlessly combined with state-of-the-art approaches for Class-Incremental Learning.

The first approach, Balanced Softmax for Class-Incremental Learning, explores the balance between old and new class in the classification loss. Rehearsal memory is a simple yet effective method to mitigate catastrophic forgetting which stores few exemplars of previously encountered classes and replays them while learning the new classes. However, due to the limited size of the rehearsal memory, the training dataset of each incremental step is highly imbalanced: containing few samples for the numerous old classes and a lot of samples for the few new classes. This results in a model biased toward the new classes. To address this issue, the author proposes to replace the standard Softmax by the Balanced Softmax in the classification loss. Experiments on competitive benchmarks show that this proposed method can improve the performance of state-of-the-art approaches while also decreasing the computational cost of their training procedure in some cases. Furthermore, the author extended this method to Exemplar-Free Class-Incremental Learning where it is not possible to store any exemplar of past classes and experimentally demonstrated that it can be used to adapt methods designed for Exemplar-Based Class-Incremental Learning to this more challenging setting.

The second approach, Memory Augmented using Diffusion Model for Class-Incremental Learning, explores the use of additional external data for balancing the training dataset itself at each incremental step for Exemplar-Based Class-Incremental Learning methods. Following the recent advances in text-to-image generative models and their wide availability, the author proposed to leverage large pre-trained text-to-image diffusion models in order to generate synthetic images of previously learned classes. Compared to concurrent approaches which rely on unlabeled real images sampled from the internet as a source of additional data, the present method can generate synthetic images belonging to the same classes as the previously encountered samples. This enables the use of those additional data samples not only in the distillation loss but also for replay in supervised losses such as the classification loss. Detailed experiments and ablation study highlight that the proposed method can be combined with various state-of-the-art approaches for Class-Incremental Learning to further improve their performances and achieves similar or even superior results compared to concurrent approaches relying on real unlabeled images as additional source of data.

Finally, the third approach, Future-Proofing Class-Incremental Learning, explores the expansion of the dataset used during the first incremental step for Exemplar-Free Class-Incremental Learning methods. Methods relying on frozen feature extractors have been recently popularized for this challenging setting due to their performances and lower computational costs. However, those methods

are highly dependent on the data used to train the feature extractor during the first incremental step and may struggle when only a limited number of classes are available. To overcome this limitation, the author proposes to predict at the time of the first step the classes that will be encountered during the following incremental steps and use synthetic samples from those classes to learn a feature extractor more suitable for the incoming tasks. The method leverages pre-trained text-to-image diffusion models in order to generate synthetic images of the possible future classes. Compared to the previous approach that leveraged pre-trained diffusion models to generate synthetic samples of past classes, the present approach aims at generating synthetic samples of future classes instead. Experiments on competitive benchmarks show that the proposed method can be used to significantly improve the performance of state-of-the-art approaches and that pre-training the model using synthetic samples of future classes achieves better performances that traditional pre-training using real samples from classes different to those in the target dataset.