# Continual Learning with Deep Generative Replay

# Abstract

Attempts to train a comprehensive artificial intelligence capable of solving multiple tasks have been impeded by a chronic problem called catastrophic forgetting.

试图训练能够解决多个任务的综合人工智能已被一种称为灾难性遗忘的长期问题所阻碍。

Although simply replaying all previous data alleviates the problem, it requires large memory and even worse, often infeasible in real world applications where the access to past data is limited.

虽然简单地重放所有先前的数据可以缓解这个问题，但它需要大的内存，甚至更糟的是，在现实世界应用中通常是不可行的，因为对过去数据的访问受限。

Inspired by the generative nature of the hippocampus as a short-term memory system in primate brain, we propose the Deep Generative Replay, a novel framework with a cooperative dual model architecture consisting of a deep generative model (“generator”) and a task solving model (“solver”).

受海马体作为灵长类大脑短期记忆系统的生成本质的启发，我们提出了深度生成重放(Deep Generative Replay)，这是一个新的框架，具有由深层生成模型（“生成器”）和任务求解组成的协作双模型体系结构。 模型（“求解器”）。

With only these two models, training data for previous tasks can easily be sampled and interleaved with those for a new task. We test our methods in several sequential learning settings involving image classification tasks.  
仅使用这两个模型，可以轻松地对先前任务的训练数据进行采样并与新任务的训练数据交错。 我们在涉及图像分类任务的几个连续学习(sequential learning)设置中测试我们的方法。

# 1.Introduction

One distinctive ability of humans and large primates is to continually learn new skills and accumulate knowledge throughout the lifetime [6]. Even in small vertebrates such as rodents, established connections between neurons seem to last more than an year [13]. Besides, primates incorporate new information and expand their cognitive abilities without seriously perturbing past memories. This flexible memory system results from a good balance between synaptic plasticity and stability [1].

人类和大型灵长类动物的一个独特能力是不断学习新技能并积累一生中的知识[6]。

即使在小型脊椎动物如啮齿动物，已建立的神经元之间的连接似乎持续了一年多[13]。

此外，灵长类动物吸收新的信息和扩展认知能力，而不会严重扰乱过去的记忆。

这种灵活的记忆系统是突触可塑性和稳定性之间良好平衡的结果[1]。

Continual learning in deep neural networks, however, suffers from a phenomenon called catastrophic forgetting [22], in which a model’s performance on previously learned tasks abruptly degrades when trained for a new task. In artificial neural networks, inputs coincide with the outputs by implicit parametric representation. Therefore training them towards a new objective can cause almost complete forgetting of former knowledge. Such problem has been a key obstacle to continual learning for deep neural network through sequential training on multiple tasks.

深度神经网络的持续学习遭受了灾难性的遗忘现象，当训练新任务时，模型对之前已训练的任务性能突然降低。

在人工智能网络中，输入与输出一致是通过隐含的参数表示。

因此，训练他们完成新任务会导致先前知识几乎全部忘记。

该问题一直是通过多任务的sequential learning的深度神经网络的持续学习的阻碍

Previous attempts to alleviate catastrophic forgetting often relied on episodic memory system that stores past data [31]. In particular, recorded examples are regularly replayed with real samples drawn from the new task, and the network parameters are jointly optimized. While a network trained in this manner performs as well as separate networks trained solely on each task [29], a major drawback of memory-based approach is that it requires large working memory to store and replay past inputs. Moreover, such data storage and replay may not be viable in some real-world situations.

之前减轻遗忘灾难的尝试通常依赖于情景记忆系统(epiosodic memory system) 储存过去的信息。

特别的是，已记录的样本有规律地重放，并与新任务的实时样本一起参与网络参数的优化。

虽然以这种方式的性能与分开网络训练单独任务的性能一样，但有一个主要缺点是这种基于记忆的方法要求大量的工作内存来存储和再现过去的输入。而且，在某种现实世界的情况下，这种存储和再现的方式是不可行的。

Notably, humans and large primates learn new knowledge even from limited experiences and still retain past memories. While several biological mechanisms contribute to this at multiple levels, the most apparent distinction between primate brains and artificial neural networks is the existence of separate, interacting memory systems [26]. The Complementary Learning Systems (CLS) theory illustrates the significance of dual memory systems involving the hippocampus and the neocortex. The hippocampal system rapidly encodes recent experiences, and the memory trace that lasts for a short period is reactivated during sleep or conscious and unconscious recall [8]. The memory is consolidated in the neocortex through the activation synchronized with multiple replays of the encoded experience [27]–a mechanism which inspired the use of experience replay [23] in training reinforcement learning agents.

值得注意的是，人类和大型灵长类动物能从有限的经历中学习新知识并且能始终保留过去的记忆。

虽然有几种生物学机制在多个层面上有所贡献，但灵长类动物大脑和人工神经网络之间最明显的区别是存在独立的，相互作用的记忆系统[26]。

互补学习系统（CLS）理论说明了涉及海马体和新皮质的双记忆系统的重要性。

海马体系统迅速编码最新经历，持续短时间的记忆痕迹在睡眠或有意识和无意识的回忆中被重新激活[8]

记忆被整合到新皮层中通过与编码的经历多次重现的同步激活完成，该机制被应用到训练强化学习(reinforcement learning)的代理的经验重现中。

Recent evidence suggests that the hippocampus is more than a simple experience replay buffer. Reactivation of the memory traces yields rather flexible outcomes. Altering the reactivation causes a defect in consolidated memory [35], while co-stimulating certain memory traces in the hippocampus creates a false memory that was never experienced [28]. These properties suggest that the hippocampus is better paralleled with a generative model than a replay buffer. Specifically, deep generative models such as deep Boltzmann machines [32] or a variational autoencoder [17] can generate high-dimensional samples that closely match observed inputs.

最近的证据表明，海马体不仅仅是一种简单的经验重放缓冲器。记忆痕迹的重新激活产生相当灵活的结果。改变重新激活会导致整合记忆中的缺陷[35]，同时一起刺激海马体中的某些记忆痕迹会产生一种从未体验过的错误记忆[28]。

这些特性表明，海马体与一个生成模型并行使用会比重现缓冲器的性能好。

具体而言，深度生成模型如deep Boltzmann machines 或 a variational autoencoder 可以生成与观察到的输入紧密匹配的高维样本。

We now propose an alternative approach to sequentially train deep neural networks without referring to past data. In our deep generative replay framework, the model retains previously acquired knowledge by the concurrent replay of generated pseudo-data. In particular, we train a deep generative model in the generative adversarial networks (GANs) framework [10] to mimic past data. Generated data are then paired with corresponding response from the past task solver to represent old tasks. Called the scholar model, the generator-solver pair can produce fake data and desired target pairs as much as needed, and when presented with a new task, these produced pairs are interleaved with new data to update the generator and solver networks. Thus, a scholar model can both learn the new task without forgetting its own knowledge and teach other models with generated input-target pairs, even when the network configuration is different.

我们现在提出一种替代方法来顺序训练深度神经网络，不参考过去的数据。

在我们的深度重放框架中，该模型保留了之前已获得的知识通过生成的伪数据的并发重放。

特别是，我们在GAN结构上训练一个深度生成模型去模拟过去数据。生成数据然后与来自过去任务解算器的响应匹配去表示旧任务，称之为学者模型(scholar model)。

生成器-求解器对可以根据需要产生伪数据和期望的目标对，并且当呈现新任务时，这些生成的对与新数据交错来更新生成器和求解器网络。

因此，学者模型可以在不忘记原有知识的条件下学习新任务并且用生成的目标对教其他模型，即使该网络配置是不同的。

As deep generative replay supported by the scholar network retains the knowledge without revisiting actual past data, this framework can be employed to various practical situation involving privacy issues. Recent advances on training generative adversarial networks suggest that the trained models can reconstruct real data distribution in a wide range of domains. Although we tested our models on image classification tasks, our model can be applied to any task as long as the trained generator reliably reproduces the input space.

由于学者网络支持的深度生成重放保留了知识而无需重新访问实际的过去数据，因此该框架可用于涉及隐私问题的各种实际情况。

关于训练GAN的最新进展表明已训练的模型可以重建各种领域的实际数据分布。

虽然我们在图像分类任务上测试了我们的模型，但我们的模型可以应用到任意任务上，只要已训练好的生成器可靠地再现输入空间。

# 2.Related Works

The term catastrophic forgetting or catastrophic interference was first introduced by McCloskey and Cohen in 1980’s [22]. They claimed that catastrophic interference is a fundamental limitation of neural networks and a downside of its high generalization ability. While the cause of catastrophic forgetting has not been studied analytically, it is known that the neural networks parameterize the internal features of inputs, and training the networks on new samples causes alteration in already established representations. Several works illustrate empirical consequences in sequential learning settings [7, 29], and provide a few primitive solutions [16, 30] such as replaying all previous data.

McCloskey和科恩在1980年代首先介绍了灾难性遗忘或灾难性干扰这一术语[22]。

他们声称，灾难性干扰是神经网络的一个基本限制，也是其高泛化能力的缺点。

虽然灾难性的原因遗忘尚未经过分析研究，但众所周知，神经网络使输入的内部特征参数化了，并且在新样本上训练网络导致已建立的表述的改变。

一些论文说明了sequential learning设置中的经验结果[7,29]，并提供一些落后的解决方案[16,30]，例如重放所有先前的数据。

## 2.1 Comparable methods

A branch of works assumes a particular situation where access to previous data is limited to the current task[12, 18, 20]. These works focus on optimizing network parameters while minimizing alterations to already consolidated weights. It is suggested that regularization methods such as dropout [33] and L2 regularization help reduce interference of new learning [12]. Furthermore, elastic weight consolidation (EWC) proposed in [18] demonstrates that protecting certain weights based on their importance to the previous tasks tempers the performance loss.

工作的一个分支假设一种特定情况，即对先前数据的访问仅限于当前任务[12,18,20]。

这些工作的重点在于优化网络参数，同时是已经合并的权重改变最小。

建议采用正则化方法（regularization method）例如dropout[33]和L2 regularization,可以减少新学习的干扰。此外，EWC提出保护某些基于先前任务的重要性的权重缓和了性能的损失。

Other attempts to sequentially train a deep neural network capable of solving multiple tasks reduce catastrophic interference by augmenting the networks with task-specific parameters. In general, layers close to inputs are shared to capture universal features, and independent output layers produce taskspecific outputs. Although separate output layers are free of interference, alteration on earlier layers still causes some performance loss on older tasks. Lowering learning rates on some parameters is also known to reduce forgetting [9]. A recently proposed method called Learning without Forgetting (LwF) [21] addresses the problem of sequential learning in image classification tasks while minimizing alteration on shared network parameters. In this framework, the network’s response to new task input prior to fine-tuning indirectly represents knowledge about old tasks and is maintained throughout the learning process.

顺序地训练可以解决多任务的深度神经网路可以通过使用特定参数扩充网络来减少灾难性干扰。

通常，接近输入的层能捕获通用特征，独立的输出层可以产生任务特定的输出。

虽然单独的输出层不受干扰，但之前层的改变仍能导致原定任务的某些性能损失。

降低某些参数的学习率可以减少遗忘。

LwF解决了图像分类任务中的sequential learning 的问题，同时最小化共享网络参数的改变。

在此框架中，网络在微调之前对新任务输入的响应能间接代表旧任务的知识，并且响应能在整个网络中维持。

## 2.2 Complementary Learning System (CLS) theory

A handful of works are devoted to designing a complementary networks architecture to alleviate catastrophic forgetting. When the training data for previous tasks are not accessible, only pseudo-inputs and pseudo-targets produced by a memory network can be fed into the task network. Called a pseudorehearsal technique, this method is claimed to maintain old input-output patterns without accessing real data [31]. When the tasks are as elementary as coupling two binary patterns, simply feeding random noises and corresponding responses suffices [2]. A more recent work proposes an architecture that resembles the structure of the hippocampus to facilitate continual learning for more complex data such as small binary pixel images [15]. However, none of them demonstrates scalability to high-dimensional inputs similar to those appear in real world due to the difficulty of generating meaningful high-dimensional pseudoinputs without further supervision.

一些工作致力于设计一个互补的网络结构去减轻灾难性遗忘。

当不能访问先前任务的训练数据时，只能由记忆网络产生的伪输入和伪目标可以被反馈到任务网络中，称为伪重造技术（pseudorehearsal technique）。

这种方法据称可以维持旧的输入输出模式，而不需要访问实际数据。

当任务与耦合两个二进制模式一样基本时，简单地输入随机噪声和相应的响应就足够了[2].

最近的一项工作提出了一个类似于海马体结构的体系结构，用来促进用于更多复杂数据的continual learning，如二进制像素图像[15]。

然而，没有人提出与现实世界类似的高维输入的可扩展性，由于难以在无进一步监督的情况下产生有意义的高维伪输入。

Our generative replay framework differs from aforementioned pseudorehearsal techniques in that the fake inputs are generated from learned past input distribution. Generative replay has several advantages over other approaches because the network is jointly optimized using an ensemble of generated past data and real current data. The performance is therefore equivalent to joint training on accumulated real data as long as the generator recovers the input distribution. The idea of generative replay also appears in Mocanu et al. [24], in which they trained Restricted Boltzmann Machine to recover past input distribution.

我们的生成重放框架与前面提到的伪理论技术不同，伪输入是从已学习的过去的输入分布生成的。生成重现(Generative replay)与其他方法相比有很多优点，因为网络是通过使用生成的过去数据和实时的当前数据的集合联合优化的。因此，性能相当于联合训练累积的实时数据，只要生成器能恢复输入分布。生成重现的想法也出现在Mocanu等人的研究中，他们训练Restricted Boltzmann Machine来恢复过去的输入分布。

## 2.3Deep Generative Models

Generative model refers to any model that generates observable samples. Specifically, we consider deep generative models based on deep neural networks that maximize the likelihood of generated samples being in given real distribution [11]. Some deep generative models such as variational autoencoders [17] and the GANs [10] are able to mimic complex samples like images.

生成模型是指生成可观察样本的任何模型。

具体来说，我们考虑基于深度神经网络的深度生成模型，这些模型使生成样本的likelihood在给定的实际分布上最大化。

一些深度生成模型如variational autoencoders,GANs 都能够模拟复杂样本如图像。

The GANs framework defines a zero-sum game between a generator G and a discriminator D. While

the discriminator learns to distinguish between the generated samples from real samples by comparing two data distributions, the generator learns to mimic the real distribution as closely as possible. The objective of two networks is thereby defined as:

GAN的框架定义了G和D之间的零和游戏。

当D学习了通过比较两个数据分布来区分生成样本和实际样本，G学会尽可能地模拟真实分布，因此两个网络的目标定义为…

# 3.Generative Replay

# 4.Experiments

In this section, we show the applicability of generative replay framework on various sequential

learning settings. Generative replay based on a trained scholar network is superior to other continual learning approaches in that the quality of the generative model is the only constraint of the task performance. In other words, training the networks with generative replay is equivalent to joint training on entire data when the generative model is optimal. To draw the best possible result, we used WGAN-GP [14] technique in training the generator.

本节中，我们展示了生成重现框架在多种sequential learning设置上的适用性。

基于已训练的学者网络的生成重现优于其他continual learning 刚发，这些方法中生成模型的质量之受任务性能的限制。

换句话说，用生成再现训练网络相当于联合训练整个数据，当生成模型是最佳的时候。

为了得到最佳的结果，我们用了WGAN-GP technique去训练G。

As a base experiment, we test if generative replay enables sequential learning while compromising

performance on neither the old tasks nor a new task. In section 4.1, we sequentially train the networks on independent tasks to examine the extent of forgetting. In section 4.2, we train the networks on two different yet related domains. We demonstrate that generative replay not only enables continual learning on our design of the scholar network but also compatible with other known structures. In section 4.3, we show that our scholar network can gather knowledge from different tasks to perform a meta-task, by training the network on disjoint subsets of training data.

作为一个基础实验，我们测试生成重放是否能使连续学习成为可能，当性能既不在旧任务也不在新任务上妥协时。

4.1节中，我们依次对独立任务训练网络，来检查遗忘的程度。

4.2节中，我们在两个不同但相关的域上训练网络。我们证明了重放再现不仅能使学者网络的连续学习成为可能，同时也能兼容其他已知的结构。

4.3节中，我们展示了学者网络可以从不同任务中收集知识来执行一个meta-task,通过训练该网络训练数据的不相交子集。

We compare the performance of the solver trained with variants of replay methods. Our model with generative replay is denoted in the figure as GR. We specify the upper bound by assuming a situation when the generator is perfect. Therefore, we replayed actual past data paired with the predicted targets from the old solver network. We denote this case as ER for exact replay. We also consider the opposite case when the generated samples do not resemble the real distribution at all. Such case is denoted as Noise. A baseline of naively trained solver network is denoted as None. We use the same notation throughout this section.

我们比较不同S的性能，这些S由不同的再现方法训练而成。我们的模型简写为GR。

我们通过假设G是完美的来指定上限。

因此，我们重放与预测目标配对(来源于旧的solver网络)的实际过去数据。称为ER。

我们也考虑了相反的情况，当生成样本与真实分布完全不像。该情况下称为Noise。

已训练的S网络的基线称为None。

## 4.1 Learning independent tasks

The most common experimental formulation used in continual learning literature [34, 18] is a simple image classification problem where the inputs are images from MNIST handwritten digit database[19], but pixel values of inputs are shuffled by a random permutation sequence unique to each task. The solver has to classify permuted inputs into the original classes. Since the most, if not all pixels are switched between the tasks, the tasks are technically independent from each other, being a good measure of memory retention strength of a network.

最常见的实验基础是简单的图像分类问题（Mnist），但输入的像素值是随机组合的序列。

S必须将组合的输入分类至原始类。

## 4.2 Learning new domains

Training independent tasks on the same network is inefficient because no information is to be shared. We thus demonstrate the merit of our model in more reasonable settings where the model benefits from solving multiple tasks.

在同一网络上训练独立任务是无效的，因为无任何信息共享，因此我们展示更多合理设置的优点。

A model operating in multiple domains has several advantages over a model that only works in a

single domain. First, the knowledge of one domain can help better and faster understanding of other domains if the domains are not completely independent. Second, generalization over multiple domains may result in more universal knowledge that is applicable to unseen domains. Such phenomenon is also observed in infants learning to categorize objects [3, 4]. Encountering similar but diverse objects, young children can infer the properties shared within the category, and can make a guess of which category that the new object may belong to.

在多个域中运行的模型比仅在一个域中运行的模型相比具有更多的优点。

首先，一个领域的知识可以帮助更好更快地理解其他领域，如果这些领域不是完全的独立。

其次，多领域的通用性会产生更普遍的知识，这些可以应用在未知领域。这个现象可以在婴儿学习过程中观察到。

当遇到相似但不同的对象，小孩可以推断出该类别共享的特征，并且猜测新对象可能属于哪一类。

Generative replay is compatible with other continual learning models as well. For instance, Learning without Forgetting (LwF), which replays current task inputs to revoke past knowledge, can be augmented with generative models that produce samples similar to former task inputs. Because LwF requires the context information of which task is being performed to use task-specific output layers, we tested the performance separately on each task. Note that our scholar model with generative replay does not need the task context.

生成重放可与其他连续学习模型兼容。例如，LwF，重放当前任务输入去撤销过去知识，可以增强模型产生与之前任务输入相似的样本。

由于LwF需要任务被执行以使用特定于任务的输出层的内容信息，所以我们独立测试每一层的性能。

请注意，我们的生成重放学者模型不需要任务内容。

In Figure 5, we compare the performance of LwF algorithm with a variant LwF-GR, where the task-specific generated inputs are fed to maintain older network responses. We used the same training regime as proposed in the original literature, namely warming up the new network head for some amount of the time and then fine tuning the whole network. The solver trained with original LwF algorithm loses performance on the first task when fine-tuning begins, due to alteration to shared network (green). However, with generative replay, the network maintains most of the past knowledge (orange).

图5中，我们比较了Lwf算法和LwF-GR算法的性能，在LwF-GR中，特定于任务的输入被反馈到旧的网络响应。

我们使用相同的训练系统，也就是…

当微调开始时，原始LwF的S损失了第一个任务的性能，由于共享网络的改变（绿色线）。

然而，使用生成重放，网络保持了大部分的过去知识。（橙色线）

## 4.3 Learning new classes

To illustrate that generative replay can recollect the past knowledge even when the inputs and targets are highly biased between the tasks, we propose a new experiment in which the network is sequentially trained on disjoint data. In particular, we assume a situation where the agent can access examples of only a few classes at a time. The agent eventually has to correctly classify examples from all classes after being sequentially trained on mutually exclusive subsets of classes. We tested the networks on MNIST handwritten digit database.

为了说明生成重放可以重新收集过去知识，即使输入和目标在任务之间高度偏向，我们提出了一个新的实验，网络依次在不相关数据集上被训练。

特别是，我们假设一种情况，代理依次只能获取一些类别的样本。代理在依次在相互排斥的子集上训练之后，最终必须正确分类所有类别的样本。

我们测试该网络使用MNIST。

Note that training the artificial neural networks independently on classes is difficult in standard settings, as the network responses may change to match the new target distribution. Hence replaying inputs and outputs that represent former input and target distributions is necessary to train a balanced network. We thus compare the variants described earlier in this section from the perspective of whether the input and target distributions of cumulative real data is recovered. For ER and GR models, both the input and target distributions represent cumulative distribution. Noise model maintains cumulative target distributions, but the input distribution only mirrors current distribution. None model has current distribution for both.

注意，独立训练人工智能网络的困难在于标准设置，因为网络响应会改变去匹配新的目标分布。因此为了训练一个均衡的网络，代表前一输入和目标分布的重放输入和输出是必要的。

我们因此比较变体（本章前面所描述的），从是否能恢复累积实时数据的输入和目标分布这个角度讨论。

对于ER和GR这些模型，两者的输入和目标分布都能表示累积分布。

Noise模型保持累积目标分布，但是输入分布只能反映当前的分布。

None可表示输入和目标的当前分布。

# 5.Discussion