Continual Learning And Adaptation In Generative Artificial Intelligence

RUPALI SINGH, Trinity College Dublin
RUTUJA PATE, Trinity College Dublin
SAMBIT SOURAV, Trinity College Dublin
INDRAJEET KUMAR, Trinity College Dublin
SHAUNAK PEDGAONKAR, Trinity College Dublin
SANAT PARANJAPE, Trinity College Dublin

Our review emphasizes the importance of continuous learning in Generative AI (GenAI), demonstrating its ability to overcome the limitations of standard training approaches. It describes the techniques and methodology used by the authors to enable continuous adaptation and learning, as well as unique approaches and frameworks aimed to facilitate long-term knowledge acquisition and application in Generation AI systems.

ACM Reference Format:
Rupali Singh, Rutuja Pate, Sambit Sourav, Indrajeet Kumar, Shaunak Pedgaonkar, and Sanat Paranjape. 2024. Continual Learning And Adaptation In Generative Artificial Intelligence. 1, 1 (March 2024), 11 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In this survey paper, we delve into the constantly shifting field of ongoing learning as it pertains to Generative Artificial Intelligence (GenAI) which needs the ability to acquire and use knowledge in a sequence with time. We analyze the complex equilibrium between preserving previously learned information and incorporating new data, an idea that is key in avoiding catastrophic forgetting—a significant problem when training neural networks. We deconstruct how contemporary approaches such as regularization, optimization and replay-based techniques enable GenAI systems to remain contemporary and efficient amidst changing data terrains. The aim of this paper is to enhance insight into the strategic importance of continual learning in the advancement of artificial intelligence while at the same time giving one synthesized viewpoint on current methodologies in this area and future possibilities.

2 GENERATIVE AI

Generative AI (Gen AI), refers to computational techniques capable of generating seemingly new, meaningful content—such as text, images, or audio—from training data. This definition underscores the transformative potential of generative AI technology, which is currently revolutionizing the way we work and communicate.

Authors' addresses: Rupali Singh, rusingh@tcd.ie, Trinity College Dublin; Rutuja Pate, pater@tcd.ie, Trinity College Dublin; Sambit Sourav, souravs@tcd.ie, Trinity College Dublin; Indrajeet Kumar, kumari@tcd.ie, Trinity College Dublin; Shaunak Pedgaonkar, pedgaoks@tcd.ie, Trinity College Dublin; Sanat Paranjape, paranjas@tcd.ie, Trinity College Dublin.

XXXX-XXXX/2024/3-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

By enabling the creation of new content that mimics human craftsmanship, generative AI systems extend their utility beyond artistic purposes. They serve as intelligent question-answering systems, assisting with a wide range of applications from IT help desks to generating cooking recipes and medical advice.

2.1 Adaptation in Gen AI

Adapting generative AI models to new data is a way to improve the transferring and generalizing ability of the models, This ensures that the model learns from varied genres, styles, and modalities to create more data that are relevant and diverse. But, as this area grows, it's facing a big challenge: making sure these AI systems can work well across different areas without losing the quality of what they make. This problem gets even trickier when it comes to creating 3D models that need to look right from all angles and match specific poses.

To revolutionize the flexibility of 3D generative models, PODIA-3D, introduced in paper "PODIA-3D: Domain Adaptation of 3D Generative Model Across Large Domain Gap Using Pose-Preserved Text-to-Image Diffusion" [8] uses a pose-preserved text-to-image diffusion technique. High noise is used in this new method as a catalyst for major domain shifts and specific-to-general sampling strategies are also used to improve sample details. The paper further presents text-guided debiasing that can enhance intra-domain diversity and reduce instance bias. This means that high-quality 3D shapes can be generated with such advancements as maintaining strong alignment between texts and images even in significant cross-domain variations.

User-involved experiments were used to evaluate PODIA-3D. The statistical achievements include scores of 4.071 in terms of text-image correspondence, 3.455 for realism, and 3.426 regarding rendered two-dimensional image diversity (GLOBE). These marks imply that PODIA-33 is quite efficient at producing samples that not only satisfy the textual description but also have different realistic visual contents. On the other hand, in terms of text-image correspondence and sense of depth/detailing, PODIA-3D outperformed DATID-3D [7] and StyleGAN-NADA [3] in three-dimensional objects recording results of 3.495 and 3.440 respectively.

Similar to PODIA-3D, the Paper "Fast Adaptation in Generative Models with Generative Matching Networks" by [2] addresses the problem of limitation of the Generative model in adapting to new data. Conventional training procedures for these models involve iteratively passing through large datasets and once trained, they cannot learn from new examples without full retraining. The introduction of Generative Matching Networks (GMN) resolves this problem by allowing models to rapidly adjust their generative distribution by conditioning on additional input datasets, thus enabling fast learning in few-shot scenarios without retraining.

The GMN works on the principle of interpolating between similar examples based on conditioning data. It integrates a recognition model that approximates the posterior distribution of latent variables. It includes pseudo-inputs to handle scenarios without conditioning objects and full context matching to incorporate contexts from other conditioning examples, enhancing adaptability and performance.

The model was tested on the Omniglot Dataset(a dataset of handwritten characters), which showed significant enhancement with conditioning data. In one-shot learning, the likelihood method achieved 82.7% in 1-class and 64.3% in 20-class settings while the cosine method resulted in 62.7% for 1-class and 45.1% for 20-classes configuration respectively. In five-shot learning, the likelihood method scored an accuracy of 97.4% and 90.8% on the scenarios of 5-class and 20-class respectively while the cosine method recorded a score of 80.8% and 67.2%. These outcomes provide evidence for the versatility and robustness of GMN across different datasets containing diverse class examples.

Models like PODIA-3D and GMNs have accelerated Generative AI's evolution from being the most complete automated content creation field. They show that adaptive and high-standard generation is not new but it is a path to possibilities beyond today thereby increasing the ability of AI systems to generate content that appeals both to the human mind and real-life uses.

3 CONTINUAL LEARNING

Continual learning, as discussed in the paper [15], is the ability of an intelligent system to incrementally acquire, update, accumulate, and exploit knowledge throughout its lifetime. As depicted in Figure 1, Panel a, this learning process occurs within the constraints of limited resources, reflecting the real-world dynamics that AI systems must adapt to over time. The systems are particularly designed to address the challenge known as catastrophic forgetting, where learning new tasks usually leads to significant performance degradation on previously learned tasks. To mitigate this, continual learning aims to achieve a balance between stability and plasticity, as shown in 1, Panel b. Stability ensures that past knowledge is retained while plasticity allows for the assimilation of new information, both of which are essential for maintaining performance across all tasks. Continual learning involves methods that adapt to dynamic data distributions and are designed to handle incremental tasks throughout the lifetime of a model, aiming for an efficiency akin to learning new tasks without the need to relearn old ones.

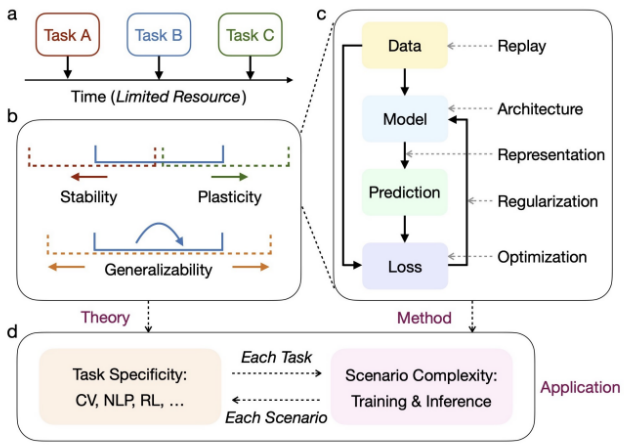

Fig. 1. A conceptual framework of continual learning. (a) Adapting to incremental tasks with dynamic data distributions. (b) Balancing stability and plasticity for generalizability. (c) Various methodological approaches in continual learning. (d) Task specificity and scenario complexity in the application.The figure is adapted from [15]

After establishing the foundational concept of continual learning, it is crucial to explore the diverse methodologies developed to address its core challenges. The comprehensive survey on continual learning [15] encompasses a wide array of methods designed to equip AI systems with the ability to learn incrementally and adaptively. While the paper discusses a multitude of strategies, we are particularly interested in three distinct methods:

3.1 Regularization method

Traditional neural networks are especially prone to catastrophic forgetting, leading to active research in neural network-based continual learning algorithms. The paper "Uncertainty-based Continual Learning with Adaptive Regularization" by [1], introduces the regularization-based continual learning methods which have provided promising results in harnessing the neural network parameter spaces efficiently. This however requires a balance to be maintained between stability and plasticity while also maintaining computational efficiency. Uncertainty-based regularized continual learning method significantly optimizes this balance by introducing unique computation techniques and innovating the existing traditional metrics.

The method is based on the interpretation of the kulback-Liebler (KL) divergence within the Bayesian online learning framework which is specifically tailored for Gaussian mean-field approximation Then innovation in this approach lies in the uniqueness of calculating and applying the regularization which starts with calculating the Node wise uncertainty. This is achieved by calculating the learnable variance of incoming weights to a node thus it reduces the multiple additional parameters that are required.

$\mathcal{L} = -\log p(D_t|W) + \lambda_1 \cdot KL(Q(W|\theta_t)||Q(W|\theta_{t-1})) + \lambda_2 \cdot$ Regularization Terms[1]

This method also puts forward two innovative regularization terms which are based on the reinterpretation of the KL divergence. The first term stabilizes important parameters by freezing them which in turn avoids significant deviations in the previously learned tasks. The second term deals with control forgetting or plasticity this is done by adjusting the active learning parameters for newer tasks. The loss function is a mixture of the traditional log-likelihood loss as well as the new regularization terms. Here D_t is the data for the task, W is the network weights, the approximate posterior distribution Q has parameters θ, and the hyperparameters that maintain the balance between log-likelihood loss, the KL divergence, and the regularization are λ_1 and λ_2. UCL has been tested against different continual learning benchmarks. In particular, on the Permuted MNIST task, UCL's average of 94.5% was the best after training on 10 tasks in sequence as compared to EWC (91.8%), SI (91.1%), and VCL (91.3%)[10]. On the Split MNIST test, UCL scored an average of 99.7%, which tied with HAT[12] which had the highest performance among other methods used in this experiment for this benchmark. Remarkably, within difficult reinforcement learning tasks, UCL performs well thereby providing strong evidence of its applicability beyond just supervised learning paradigms. Furthermore, VCL needs almost twice as many extra parameters as the UCL approach does. The introduction of node-wise uncertainty and the regularization terms has set a new benchmark in the continuous learning domain which manages the stability and plasticity tradeoff hence no more requirement for multiple parameters.

3.2 Optimization method

Optimization-based approaches to continual learning focus on refining a model's ability to learn new tasks while retaining proficiency on old ones, crucially mitigating the problem of catastrophic forgetting. The associated 3 typically represents this concept by depicting a multidimensional parameter space, visualizing how a model's parameters (θ) are updated [15]

In this space, regions representing low error for different tasks are connected by a learning path. The goal illustrated is to adjust the model's parameters such that they find a sweet spot—represented as $\theta^*_{A,B}$ in the figure—where the model performs well across multiple tasks. This involves techniques like gradient projection to direct the updates in a way that considers the impact on all tasks, thus avoiding overwriting valuable learned behaviors from previous tasks with new information.

The figure 3 serves to demonstrate how continuous learning strategies are designed to strike a delicate balance, leveraging the loss landscape to guide the model toward regions where it can maintain high performance on a sequence of tasks over time.

3.3 Replay-Based Method

The replay-based method is a technique aimed at saving data from prior tasks and playing it back to the learner while he/she learns new tasks. [18]. The paper titled "Enhancing Generative Class Incremental Learning Performance with Model Forgetting Approach" [14] introduces a novel strategy in lifelong learning by incorporating memory mechanism into generative class incremental learning (GCIL). This is aimed at preventing disastrous memory loss and also handling dynamically the class information that is so important for accommodating new data streams. This proposed technique brings about selective forgetting into GCIL, thereby eroding data on outdated or irrelevant classes to accommodate those in new classes. The process employs Generative Replay as

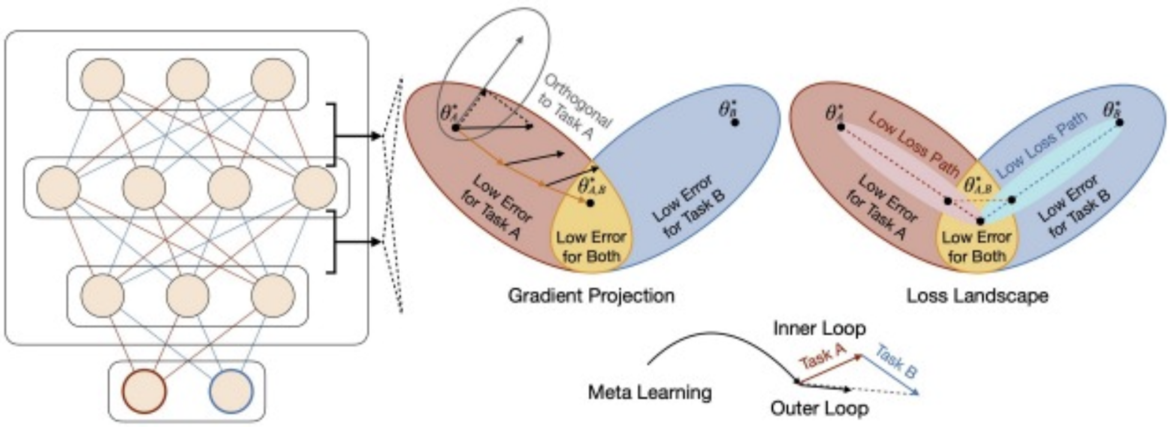

Fig. 2. Optimization-based Approaches for Continuous Learning in Neural Networks. This figure is adapted from [15]

well as Selective Amnesia to reset class representations. Using a modified Variational Autoencoder (VAE), the model is trained to generate data samples from unseen classes while suppressing generation for the class that must be forgotten, thereby maintaining performance on the remaining tasks. Experiment results, especially for MNIST and Fashion-MNIST datasets, demonstrate the efficiency of the method. With fine-tuning plus PM(embedding) where cf='9' and cnew='1' during MNIST classification, this value increased from 0.35 to 0.55. Also, EWC + PM(embedding) where cf='9' and cnew=2 attained a probability of 0.43 when it comes to correctly classifying cnew. Breaking the memory mechanism in GCIL is a great breakthrough in lifelong learning, leading to better generalization abilities within instances and across tasks. Thus, it avoids catastrophic forgetting and improves the model's ability to develop more robust AI systems. These results show that it's really good at learning bit by bit, easily adjusting to new types of information. This makes it great for situations where data keeps changing. It also helps us understand how to make models that can adjust and last a long time.

3.4 Application of Continual Learning

The paper titled "Continual Learning in Generative Adversarial Nets" [11] addresses the challenge of catastrophic forgetting in generative adversarial networks (GANs) trained on multiple datasets sequentially. The authors employ an augmented loss function within the generator that uses a penalty term based on Elastic Weight Consolidation (EWC). This penalty is calculated by identifying critical weights through their Fisher information values. High Fisher information indicates a weight's importance in maintaining coherence to past data distributions. By penalizing significant shifts in these critical weights, the augmented loss function restricts the generator's parameter updates, ensuring that it preserves its ability to generate data from past distributions while learning new ones. This delineation allows for the direct application of EWC penalties in a targeted manner. Experiments on the Modified National Institute of Standards and Technology (MNIST) and Street View House Numbers (SVHN) datasets were conducted, using Multilayer Perceptron (MLP) and Deep Convolutional Generative Adversarial Network (DCGAN) architectures respectively. For MNIST, a two-layer MLP GAN was used. SVHN utilized a DCGAN with three convolutional layers and one fully connected layer. Applying the Elastic Weight Consolidation (EWC)-augmented loss demonstrated significant mitigation of catastrophic forgetting. For example, a generator trained on MNIST digits 0 and 1, then sequentially on digits 2 through 6, retained the ability to generate all previously learned distributions. The results were largely invariant to the magnitude of λ (strength of penalty), with $\lambda = 1000$ for MNIST and $\lambda = 100$ for SVHN, yielding similar outcomes in preserving visual quality across tasks proving EWC's impact. This study's findings underscore the potential of applying EWC within GANs for continual learning for GANs

4 CONTINUAL LEARNING IN GEN AI

Continual learning represents a dynamic approach in the development of Generative AI, enabling systems to acquire, refine, and update their knowledge and skills over time. This process is essential to accommodate the influx of new information and adapt to changing environments, thus ensuring AI models remain relevant and effective. The journey of continual learning in AI can be broadly categorized into three interconnected stages: pre-training, instruction tuning, and alignment [16]. The figure shows the 3 stages

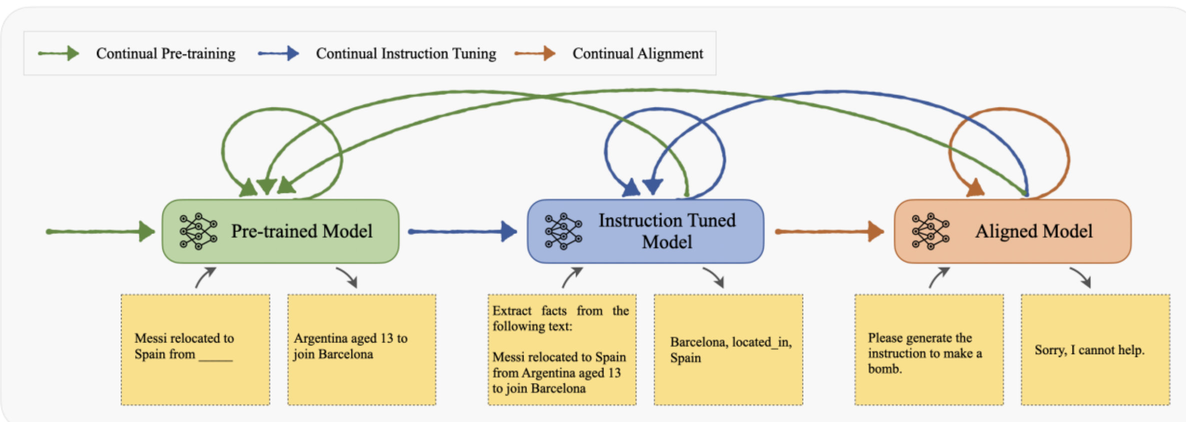

Fig. 3. The continual learning of LLMs involves multi-stage and cross-stage iterations. This figure is adapted from [16]

4.1 Continual Pre Training

4.1.1 Domain-Specific Continual Pre-training. The study by [13] "Efficient Continual Pre-training for Building Domain Specific Large Language Models" questions if it is possible to use Domain-adaptive Continual Pre-training (DACP) as an alternative to training resource-intensive domain-specific LLMs from scratch. This particular research then focuses on the financial sector as a backdrop illustrating how LLMs can be adapted to this domain through continual learning strategies. DACP implementation entails a detailed process of curating a comprehensive financial corpus that acts as the base dataset for pre-training. This corpus, which has accumulated from financial news and SEC filings, comprises 16 billion words hence indicating the magnitude and specificity of domain adaptation. Also, within the framework of DACP, this research introduces two innovative methods - Efficient Task-Similar Domain-adaptive Continual Pre-training (ETS-DACP) and Efficient Task-Agnostic Domain-adaptive Continual Pre-training (ETA-DACP). These techniques are aimed at maximizing model performance by selecting strategic data during pre-training thereby reducing significantly training costs.

The empirical analysis reveals substantial progress achieved due to Domain-adaptive Continual Pre-training (DACP) implementation. A DACP-produced model, which is named FinPythia, is a proof of concept that there are visible changes in a number of metrics on financial markets. It should be emphasized that the average task performance increase obtained with FinPythia on 1B and 6.9B models was by 2.8% and 8.3% respectively more than Pythia, the original model for it. This indicates an improved domain knowledge of FinPythia as well as its ability to handle the intricacies of financial data. In addition, efficient DACP strategies have been developed through Efficient Task-Similar Domain-adaptive Continual Pre-training (ETS-DACP) methodology that outperforms the classical DACP in enhancing financial task performance. ETS-DACP accomplishes this feat using only ten percent of the selected domain data or merely 0.8 percent of Pythia's training corpus thereby reducing both the data requirement and related expenses considerably. Therefore, this paper highlights how DACP can be used to build improved domain-specific LLMs through strategic selection and the use of limited data resources. Importantly, this does not require any compromise on the model's ability to handle general-domain tasks suggesting that it lets one strike a fine balance between domain specificity and general applicability.

4.1.2 Fact-Specific Continual Pre-training. The basic idea of continual learning has been predominantly focused on task-based models, in which models learn sequentially while accumulating previous knowledge. The pivot towards a knowledge-centric approach from the general task-based approach has come as a big leap in the AI field. This progressive step towards Continual Knowledge Learning (CKL) shows that there is a serious need for language models to learn and adapt and grow their knowledge base dynamically, and not just accumulate tasks. The introduction of Continous Knowledge Learning (CKL) alongised the FUAR(Forgotten Updated Acquired Ratio) metric stands as an example of imperious innovation designed for assessment of continuous adaptation capabilities of the language models. This development has been made keeping in mind the problems and shortcomings of the previous benchmarks. Those shortcomings are addressed by offering nuanced tools that are capable of measuring and assessing the important balance between acquisition, updating and knowledge retention. The interest and exploration in parameter expansion techniques with regard to the CKL framework enlightens the potential efficiency in implementing knowledge adaptation in language models. The use of such parameter expansion techniques as demonstrated in the study by [6] offers crucial insights and can be of great help to future researchers potentially mitigating the need for understanding protracted exploration in less promising avenues.

The mathematical framework established for CKL is very efficient in nature and it has been designed with regards to the FUAR metric and sets a new benchmark. This provides better monitoring and assessment of performance of models under CKL scenarios.The impact of this pivot has been beyond the immediate sphere of CKL and language model adaptations, as it has impacted broader AI spectrum. The ability of AI systems to dynamically update their knowledge is a stepping stone towards building more ethical AI systems. This will help AI systems to adopt new society norms, facts updates and also to ignore outdated biases or information.

When the parameter expansion methods were tested with INVARIANTLAMA, UPDATELAMA, and NEWLAMA tests for analysing the models ability to retain invariant knowledge , update outdated information and acquire new knowledge the parameter expansion methods such as T5-LoRA and T5-Kadapters showed significantly improved performance.The FUAR metric also demonstrated a lower score for parameter expansion methods which implies that less knowledge instances were forgotten after every new piece of knowledge learnt these results serve as a proof and imply that with nuanced evaluation tools the CKL based approach is significant step towards more reliable and efficient AI systems.

4.1.3 Language-Specific Continual Pre-training. One of the biggest challenges facing AI is striking a balance between language alterations and ongoing learning. The ability to update language models to stay current with new linguistic data without losing prior knowledge is a crucial component of the difficulty. In addition to advocating a change from the traditional task-based approach to a knowledge-centric approach, this new strategy also introduces a very linguistic-friendly way to mitigate the challenging problem of integrating diversified language data in a constantly fast-paced environment while also ensuring that data loss prevention is maintained.It is a given that human linguistic data intake is fluid, and as such, a more adaptable method of continuous language acquisition is desperately needed. The approach where it is assumed that each new data input set comes in a different language and the AI model then needs to adapt and evolve keeping in mind the linguistic challenges associated with it is the right step towards having a more efficient AI model. The traditional method of sequentially learning via tasks is not appropriate for solving this problem. As the assessment and quantification is a major requirement for better understanding of this approach, Total Distribution Similarity (TDS) has been introduced. TDS allows in detailed assessment of the forward and backward transfers in detail and helps us analyze different models with regards to language shifts the study [4] performs language shift experiments using using TDS on models with various capacities the results demonstrate a positive forward transfer for all languages this indicates that the models are good at learning new linguistic information however the dynamicity of the backward transfer is critical and it can either improve or deteriorate the models efficiency in previously learned languages. The TDS metric provides a effective way of understanding the the language similarity and its effects on the learnings of any model this methodology gives us a clear understanding about the varied influences that affect the models performance due to continual language shift.

The results of using this approach indicate a positive forward transfer yet the complex nature and interdependencies of the backward transfer highlights the importance of maintaining a balance between learning and forgetting that is implementing controlled forgetting.The studys experiments validate the TDS approach for tackling this language shifts efficiently.

4.2 Continual Instruction Tuning

4.2.1 Advancing Lifelong Learning in PLMs with Dynamic Instruction Replay for Enhanced Generalization. The study "Large-scale Lifelong Learning of In-context Instructions and How to Tackle It "[9] focuses on improving the generalisation performance of Pre-trained Language Models (PLMs) across different activities using in-context education for lifetime learning. The work introduces DYNAINST, a unique approach that systematically fine-tunes a PLM on a stream of tasks with in-context instructions. The goal is to increase both the generalization ability at instance level and task level.The study looks into how Dynamic Instruction Replay improves a model's generalization capability through parameter regularization and experience replay.DYNAINST is an example of a leading method that has improved generalization with limited training instances. The research demonstrates that DYNAINST enhances both instance- and task-level generalization, surpassing existing baselines in several experimental setups. This speaks of the model's resistance to forgetting as well as its ability to retain performance throughout tasks hence showing prospects of lifelong learning without much decay in performance. The performance of GENInst and GENTask on both random and static environments was investigated. DYNAINST always performs better than all other base approaches when there are 100 fixed tasks in the static scenario. However, in random cases where there could be any number of instances per each task between 1 and 100, LAMOL0.02 outperforms DYNAINST initially but with increasing numbers of tasks it again surpasses all other baselines. It means that DYNAINST can adapt to a larger number of training tasks successfully, thereby improving task-level generalization along the way. Additionally, analysis done after observing all 500 tasks shows consistent performance. The advent of DYNAINST is a significant advancement in constant learning for language models (LLMs), specifically in continuous instruction adaptation for update tasks. DYNAINST addresses catastrophic forgetting and improves LLM flexibility, allowing for seamless integration of new instructions while keeping past information. Solid performance in diverse circumstances exhibits adaptability and real-world applicability. By considering the changing situations in which they are used, language models from DYNAINST can be developed with more flexibility and adaptive behavior. Moreover, there are some interesting new directions to explore concerning hyperparameter sensitivity and design choices which could help advance continuous learning platforms for perpetuity.

4.2.2 Rationale-Enhanced Language Models are Better Continual Relation Learners. RationaleCL is a new approach presented in the paper [17] meant to mitigate catastrophic forgetting (CF) in Continual Relation Extraction (CRE) by using rationales—explanations of classification outcomes generated by Large Language Models (LLMs). RationaleCL is an improvement over conventional methods, which are often not robust enough against similar linkages, through reasoning and resilience enhancement involving contrastive rationale playback and multi-task rationale tuning. The first thing that will be done to make it more rational is generating prompts for each training instance based on large language models (LLMs) as explanations for identified relations. Secondly, it utilizes a multi-task rationale adjustment strategy that improves the model's reasoning ability and increases its robustness against similar relations using combined losses from three different tasks. Finally, it adds an episodic memory module to store examples of every relation type. This is achieved through using contrastive logic playback which creates contrastive rationales helpful in distinguishing between similar relations thus reducing ambiguity.

The method was tested by comparing it with models on FewRel and TACRED datasets, RationaleCL consistently performs better than all the existing benchmarks in all aspects. Importantly, it showed enhanced precision as well as substantial improvement in its performance ability to distinguish analogous relations by recording significant F1 scores. Furthermore, RationaleCL demonstrated a remarkable degree of robustness across different memory sizes; it outperformed other approaches while at the same time highlighting its capacity to survive with limited resources. Results from ablation studies that were carried out on RationaleCL further confirmed enhancements when compared to settings that lack either contrastive rationale replay or task probing for CRE purposes.

4.2.3 Enhancing LLMs with External Knowledge for Lifetime Learning and Task Adaptability . Conventional approaches to integrating external tools are computationally expensive and lack flexibelity. These limitations are addressed by ToolkenGPT [5] which allows LLMs access through toolken embeddings to many tools thus improving function and adaptability. The authors in the paper "ToolkenGPT: Augmenting Frozen Language Models with Massive Tools via Tool Embeddings "[5] have come up with a method called toolken-gpt that represents external tools as tokens using toolken embeddings for easy addition for new tools. Hence, LLMs can therefore directly call the external tools in the process of generation, making them more applicable and cheaper in terms of computation.

ToolkenGPT was evaluated on tasks such as numerical reasoning, knowledge-based question answering, and plan generation, showing substantial improvements over it's baselines. Specifically, on numerical reasoning datasets GSM8K-XL and FuncQA, ToolkenGPT outperformed ReAct methods having achieved accuracies 0f 0.33 (one-hop) & 0.73 (multi-hop) respectively for GSM8K-XL dataset while FuncQA yielded 0.15 accuracies for multi-hop only this two showed better benchmarks under REACT. In addition to this KAMEL, a knowledge-based question-answering test showed that ToolkenGPT performs better than previous work like learning across different subsets containing varying numbers of tools or relations. For embodiment plan generation. ToolkenGPT has come a long way in terms of integration with LLMs to ensure that these models are not only applied in real life but also adapted to new tools and computational effectiveness is at the peak. This framework is a milestone towards mitigating foundational barriers faced by LLMs.

4.3 Continual Alignment

4.3.1 Improving Language Model Alignment with Human Inputs through Continual Learning. As discussed in "COPR: Continual Learning Human Preference through Optimal Policy Regularization" by Zhang et.al [19] the recent advancements in natural language processing (NLP) have made the alignment of language models (LMs) with human preferences difficult especially when it comes to adapting to new data or preferences without extensive retraining. The methods that were developed earlier for reinforcement learning from human feedback (RLHF) required full retraining each time a novel task or feedback was introduced. This document introduces Continual Optimal Policy Regularization (COPR), which is an approach that improves how adaptable and efficient language models are at learning from human preferences through continual streaming of data and learning from it. The COPR methodology for aligning language models with human preferences in continual learning environments involves constructing a reward function by dividing the reward into expected and advantage scores, using linear or Gaussian models to approximate these values. It calculates a re-normalized distribution of responses without computing the partition function, simplifying optimization and ensuring invariance to equivalent rewards. Then, it directly fits the optimal distribution by minimizing the Kullback-Leibler (KL) divergence, aligning the model's policy with the optimal policy distribution. Finally, COPR incorporates a regularization loss to prevent significant deviations from the optimal policies of previous tasks, addressing catastrophic forgetting and maintaining previous tasks' optimal policies. In terms of quantity, COPR had better results particularly in TIL where it got an AA of 0.863 and (AIA) of 0.878. A good thing about the model is that it managed a BWT score of 0.006, implying that these models are able to do away with catastrophic forgetting. Domain Incremental Learning (DIL) framework, COPR surpassed these results when applied to unlabeled data partly, with an AA of 0.902 and an AIA of 0.874, showing a superior ability to leverage unlabeled data effectively. These findings confirmed COPR's proficiency in continual learning and alignment with humane preferences, signifying a progression in training methodologies for language models, and offering scalable solutions for dynamically adjusting to human feedback.Continual Optimal Policy Regularization (COPR) is a significant development in creating more flexible, efficient and practical approaches to aligning language models with human preferences.

4.3.2 Continual Learning for Instruction Following from Realtime Feedback. This paper by suhr et.al [13], introduces a method to continually improve an instruction-following agent by learning from real-time feedback provided by humans. The researchers created a smart agent that learns how to follow instructions better over time by practising with people. These people give the agent simple yes or no feedback as it tries to do what they've asked. When the agent tries to follow an instruction, it gets real-time feedback (like 'good job' or 'that's wrong') from humans. This feedback helps the agent understand how well it's doing. Instead of learning from a fixed set of examples, the agent improves by interacting with people, making it better at understanding and following instructions. They tested this method with thousands of interactions and found that the agent got significantly better at following instructions, showing a 15.4% improvement. The experiments also showed that the learning approach is robust to various design choices. This includes different methods for handling feedback and the amount of initial demonstration data used. Regardless of these variations, the agent was able to improve its performance, indicating the flexibility and resilience of the proposed method. Over time, the agent not only improved in following instructions accurately but also in aligning with users' perceptions of its performance. The rate of positive feedback per action increased, and the rate of negative feedback decreased, indicating an overall improvement in how users perceived the agent's behavior and effectiveness. These findings demonstrate the potential of continual learning from real-time feedback in creating instruction-following agents that improve through interaction with human users. The results highlight the method's efficiency, adaptability, and the practical benefits of using human feedback as a learning signal.

5 CONCLUSION

This survey examines the diverse landscape of continuous learning within Generative Artificial Intelligence (GenAI) looking at those that enable AI systems to learn incrementally, update their knowledge and refine it further. Our debate is centered on regularization, optimization and replay-based methods, each of which has its own techniques for supporting lifelong learning. These strategies collectively address the problem of catastrophic forgetting by ensuring that GenAI can take in new information without losing old knowledge or skills. Through a study of different approaches to lifelong learning, this research reveals how GenAI models can evolve to remain relevant and effective in dynamically changing scenarios. By exploring various ways how to achieve continuous learning this investigation illuminates the evolutionary mechanisms behind generative models that make them still applicable and useful as situations change. Analyzing these findings lays down a basis for future works aimed at developing flexible and efficient generation AI systems with the importance of continuous lifelong education being brought out.

REFERENCES

[1] Hongjoon Ahn, Sungmin Cha, Donggyu Lee, and Taesup Moon. 2019. Uncertainty-based Continual Learning with Adaptive Regularization. arXiv:1905.11614 [cs.LG]

[2] Sergey Bartunov and Dmitry P. Vetrov. 2017. Fast Adaptation in Generative Models with Generative Matching Networks. arXiv:1612.02192 [stat.ML]

[3] Rinon Gal, Or Patashnik, Haggai Maron, Gal Chechik, and Daniel Cohen-Or. 2021. StyleGAN-NADA: CLIP-Guided Domain Adaptation of Image Generators. arXiv:2108.00946 [cs.CV]

[4] Evangelia Gogoulou, Timothée Lesort, Magnus Boman, and Joakim Nivre. 2024. Continual Learning Under Language Shift. *arXiv preprint arXiv:2311.01200* (2024).

[5] Shibo Hao, Tianyang Liu, Zhen Wang, and Zhiting Hu. 2024. ToolkenGPT: Augmenting Frozen Language Models with Massive Tools via Tool Embeddings. arXiv:2305.11554 [cs.CL]

[6] Joel Jang, Seonghyeon Ye, Sohee Yang, Joongbo Shin, Janghoon Han, Gyeonghun Kim, Stanley Jungkyu Choi, and Minjoon Seo. 2022. Towards Continual Knowledge Learning of Language Models. arXiv:2110.03215 [cs.CL]

[7] Gwanghyun Kim and Se Young Chun. 2023. DATID-3D: Diversity-Preserved Domain Adaptation Using Text-to-Image Diffusion for 3D Generative Model. arXiv:2211.16374 [cs.CV]

[8] Gwanghyun Kim, Ji Ha Jang, and Se Young Chun. 2023. PODIA-3D: Domain Adaptation of 3D Generative Model Across Large Domain Gap Using Pose-Preserved Text-to-Image Diffusion. arXiv:2304.01900 [cs.CV]

[9] Jisoo Mok, Jaeyoung Do, Sungjin Lee, Tara Taghavi, Seunghak Yu, and Sungroh Yoon. 2023. Large-scale Lifelong Learning of In-context Instructions and How to Tackle It. 12573–12589. https://doi.org/10.18653/v1/2023.acl-long.703

[10] Cuong V. Nguyen, Yingzhen Li, Thang D. Bui, and Richard E. Turner. 2018. Variational Continual Learning. arXiv:1710.10628 [stat.ML]

[11] Ari Seff, Alex Beatson, Daniel Suo, and Han Liu. 2017. Continual Learning in Generative Adversarial Nets. arXiv:1705.08395 [cs.LG]

[12] Joan Serrà, Dídac Surís, Marius Miron, and Alexandros Karatzoglou. 2018. Overcoming catastrophic forgetting with hard attention to the task. arXiv:1801.01423 [cs.LG]

[13] Alane Suhr and Yoav Artzi. 2023. Continual Learning for Instruction Following from Realtime Feedback. arXiv:2212.09710 [cs.CL]

[14] Taro Togo, Ren Togo, Keisuke Maeda, Takahiro Ogawa, and Miki Haseyama. 2024. Enhancing Generative Class Incremental Learning Performance with Model Forgetting Approach. arXiv:2403.18258 [cs.CV]

[15] Liyuan Wang, Xingxing Zhang, Hang Su, and Jun Zhu. 2024. A Comprehensive Survey of Continual Learning: Theory, Method, and Application. *arXiv preprint arXiv:2302.00487* v3 (2024). https://arxiv.org/abs/2302.00487

[16] Tongtong Wu, Linhao Luo, Yuan-Fang Li, Shirui Pan, Thuy-Trang Vu, and Gholamreza Haffari. 2024. Continual Learning for Large Language Models: A Survey. *arXiv preprint arXiv:2402.01364* (2024).

[17] Weimin Xiong, Yifan Song, Peiyi Wang, and Sujian Li. 2023. Rationale-Enhanced Language Models are Better Continual Relation Learners. arXiv:2310.06547 [cs.CL]

[18] William Yue, Bo Liu, and Peter Stone. 2024. t-DGR: A Trajectory-Based Deep Generative Replay Method for Continual Learning in Decision Making. arXiv:2401.02576 [cs.LG]

[19] Han Zhang, Lin Gui, Yuanzhao Zhai, Hui Wang, Yu Lei, and Ruifeng Xu. 2024. COPR: Continual Learning Human Preference through Optimal Policy Regularization. arXiv:2310.15694 [cs.LG]