Incremental Deep Learning for Defect Detection in Manufacturing

Subjects: Computer Science, Artificial Intelligence Contributor: Reenu Mohandas , Mark Southern , Eoin O'Connell , Martin Hayes

Deep learning based visual cognition has greatly improved the accuracy of defect detection, reducing processing times and increasing product throughput across a variety of manufacturing use cases. There is however a continuing need for rigorous procedures to dynamically update model-based detection methods that use sequential streaming during the training phase.

deep learning incremental learning continuous learning

1. Introduction

Model-based deep learning has long been viewed as the go-to method for the detection of defects, process outliers and other faults by engineers who wish to use artificial intelligence within computer vision-based inspection, security or oversight tasks in manufacturing. The data-hungry nature of such deep learning models that have accompanied the proliferation of AI-enabled data acquisition systems means that the new information that is either captured or inferred in realtime by sensors, IoT devices, surveillance cameras and other high definition images must now be gathered and analysed in an ever smaller time window. Additionally, transfer learning and transformer networks are now providing researchers with the capacity to build pre-trained networks ^[1], which, when coupled with a deep learning framework, enable the generation of diagnostic outputs that give highly accurate real-time answers to difficult inspection questions, even in those cases where only relatively small datasets are known to exist a priori. The requirement that data need to be independently and identically distributed *(i.i.d)* across the training and test dataset has arisen with the advent of transfer learning and the availability of pre-trained models that can be 'fine-tuned' for a particular task at hand.

Classically, deep learning models were trained based on an underlying assumption that all the possible exceptions to be detected were available within the dataset a priori. In such an *offline* learning setup, sufficient numbers of static images would be collected, labelled and classified into fixed sets or categories ^[2]. The generated datasets would be then further divided into training and test subsets, where the training data would be fed into the network for sufficiently many epochs and test evaluation performed so that a high level of confidence would exist that all possible faults could be reliably detected. Such an approach to defect detection has its roots in the AdaBoost algorithm ^[3] and papers therein. Limiting factors in such an approach include that the classifier parameters are generally fixed and huge amounts of data are required, making the whole process cumbersome and resource intensive.

The pre-design phase that is invariably required for offline training is the most important differentiating factor between offline and *online* or *dynamic* approaches to detection or inspection. Online models are tuned using exception data from particular examples of interest rather than simply being restricted to a larger fixed set ^[4]. Complete re-training of existing models is an expensive task.

The way in which new classes/tasks are added to a deep learning model once it has been deployed in the field for a specific inspection task is a recurring engineering challenge for the deployment of AI in manufacturing. **Figure 1** is an illustration of process cycle without incremental learning. In a practical real-world setting, decision loops based on continuous, temporal streams of data (of which only a narrow subset maybe potentially actionable) must be considered dynamically so that categorisation, exception handling and object identification are not pre-designed or classified a priori.



Figure 1. Visual cognition engine development process cycle using deep learning without incremental learning. The process stages end with model deployment, and the lifelong learning process is incomplete without incremental learning algorithms to update the deployed model-based detection system.

2. Incremental Learning Frameworks for Deep Learning

Deep learning models are often trained using standard backpropagation, where the training corpus that is available is fed in its entirety into a model in batches. This is not a favoured approach in defect detection where new defects can arise aperiodically in a manufacturing setting. The problem is exacerbated if re-manufacturing is required. In areas where data are available as a dynamic stream, the amount of data will incrementally grow, making the storage space insufficient. Network depth is an important factor when dynamically learning complex patterns in an incremental fashion. Deep, parameter-rich models face the problem of slow convergence during the training process ^[5], but data-driven real-time inspection methods powered by deep learning models have been shown to exhibit significantly improved efficiency in manufacturing use cases ^[6].

2.1. Incremental Learning in Manufacturing

With the advent of intelligent machines, human-machine collaboration in decision making has evolved into a 'peerto-peer' interaction of cognition and consensus between a human and the machine, contrary to the 'master–slave' interaction, which existed decades ago ^[Z]. Incremental learning benefits significantly from new-found methods of collaboration based on cognitive intelligence. Human-in-the-loop hybrid intelligence systems codify required human intentions and future decisions for the downstream tasks, which cannot be directly derived by machines. Having built automated production lines and factory settings, there has been the realisation that humans are necessary to provide an effective control layer, yet humans are often considered as an external or unpredictable component in an ML loop ^[8]. This is where human-in-the-loop oversight becomes important so that operator safety mechanisms can be safely incorporated within inspection loops. When humans interact with machines in heavily automated Industry 4.0-type factory settings, data gathering, storage and distillation are required to transform data into information for modelling purposes, thereby facilitating further automated inspection and the creation of intelligent, self-healing systems ^{[9][10]}. The operation of transforming streaming data into trainable information by selection of exemplars representing new class/category needs expert human oversight. The training stage incorporates storage of processed data, methods to alleviate catastrophic forgetting and further training with new categories. Schematic representation of incremental learning operations for manufacturing environment is given in **Figure 2**.



Figure 2. Representation of incremental learning in a visual cognition engine using deep learning.

The challenge of the proportion of defective components being significantly lower than healthy parts during inspection and reliable deep learning model update with the new types of products and defects in the manufacturing setting has been considered in ^[11]. Current trends appear to focus on the concepts of model updating, retraining and online learning that are capable of updating deployed models online in real time, even when new recycled products are added into a production line. The one-stage detector framework EfficientDet, presented in ^[12] proposes a family of eight models (D0–D7) that can be used used in a particular deployment experiment depending on the availability of resources, accuracy and complexity of the data.

Intrusion detection and network traffic management systems is another sector that uses online learning for realtime identification of newer threats based on traffic flow patterns ^[13]. This idea of online learning is also crucial in a factory floor with automated systems, sensors and IoT devices. In ^[14], a restricted Boltzmann machine (RBM) and a deep belief network are used in an attempt to learn and detect new attacks online. The RBM network is proposed as an unsupervised feature extractor. Restricted Boltzmann machine networks and auto-encoder networks are known to be suitable for feature extraction problems with unlabelled data. The ability of deep auto-encoder networks to learn hierarchical features from unlabelled data is taken advantage of in incremental learning ^[15].

Active learning has been in ceaseless research in many classification systems, information retrieval and streaming data applications, in which it has been found to require more storage for the continuous update of these classifiers ^[15]. Continuous learning problems in machine learning using ensemble models were a common stream of research where new weak classifiers were trained and added to the existing ones as new data are available and outputs from these are weighted to arrive at the final classification decision. The challenges here will be the increase in the number of weak classifiers after more iterations or updates in the model.

Research in neural networks advances by drawing more interest into the concepts of incremental learning and catastrophic forgetting ^[16]. Categories of incremental learning algorithms developed in recent research are given in **Figure 3**. Among the two modes of training using gradient descend, online mode could be more appropriately used for incremental training with weight changes being computed for each instance as new training samples are added, as opposed to batch mode training where weight changes are computed over all accumulated instances. The elastic weight consolidation (EWC) algorithm proposed by the authors constrains the weight update for important tasks from previous learning, with the importance of the task being assessed from the probability distribution of data and prior probability being used to calculate conditional probability. This log probability calculated is taken as the negative of the loss function, which then is accounted as the posterior probability distribution for the entire dataset. The posterior probability of previously learned tasks hence extracted form the basis of constraining weight update for these previously learned tasks.



Figure 3. Categories in incremental learning [17][18].

2.2. Catastrophic Forgetting

Catastrophic forgetting has been a well-known concern in the field of deep learning. It is a condition that is known to arise when a neural network loses information that has been previously learned when it is subjected to retraining or a training phase is revisited due to new information gained on subsequent or downstream tasks ^{[16][19][20]} ^[21]. Shmelkov et al. describe catastrophic forgetting as "an abrupt degradation of performance on the original set of classes when the training objective is adapted to new classes" ^[22]. The reason for this degradation in performance of previously learned information is often attributed to the *stability–plasticity dilemma*. McCloskey and Cohen have observed this problem of new learning interfering with the existing trained knowledge in neural networks in their research involving arithmetic fact learning using neural networks in their 1989 study ^[23].

2.3. Stability–Plasticity Dilemma in Neural Networks

The stability–plasticity dilemma in neural networks has been studied as another direct reason for a drop in performance of previously learned tasks; stability refers to retaining the encoded previously learned knowledge in neural networks, and plasticity refers to the ability to integrate new knowledge into these neural networks ^[17]. Effects of incremental learning on large models and pre-trained ones have been studied by Dyer et al., and their finding was that compared to the randomly initialised models, large pre-trained models and transformer models are more resistant to the problem of catastrophic forgetting ^[24]. Improved performance in deep learning models has been undeniably tied to the larger size of the training dataset and deeper model size. The empirical studies have been largely conducted on language models, GPT ^[25] and BERT ^[26], which are pretrained using the large corpora of natural language text and thereby falls into category of unsupervised pretraining. As opposed to language models, image models are pre-trained in a supervised manner.

For the question of task-incremental or task-specific learning that often occurs in AI-enabled inspection, a neural network will receive sequential batches of data assigned to a specific task, the second is class-specific learning, for which the neural networks are most widely used equally in research and industry. The authors further divide task-specific incremental learning into three categories based on the storage and data usage in sequential learning: (a) replay method, (b) regularisation-based method, (c) parameter isolation method. Replay methods store samples in raw format or use generative models to generate samples. The rehearsal method is one among the replay

methods where the neural network is retrained on a subset of selected samples for new tasks, but this method can lead to over-fitting.

2.4. Methods to Alleviate Catastrophic Forgetting

2.4.1. Replay Methods

Generative models are used to generate synthetic data or stored sample data are used to rehearse the model by providing this as input during the training of a new task. Due to the reuse of stored examples of data or generated synthetic data, this learning process is prone to over-fitting, hence constrained optimisation has been proposed ^[1,7]. iCaRL ^[2,7] is one of the earliest replay-based methods developed in 2017, which uses a nearest class prototype classifier algorithm to mitigate forgetting and has been viewed as a benchmark since its introduction. iCaRL used the nearest mean of exemplars for feature representation, reducing the required number of exemplars per class for replay and feature representation from stored exemplars combined with distillation to alleviate catastrophic forgetting. REMIND ^[1,8] is another replay-based method where quantised mid-level CNN features were stored and used for replay. Encoder-based lifelong learning (EBLL) ^[28] is an autoencoder-based method to preserve previously learned features related to the old task as an effort to alleviate forgetting. Replay-based methods such as REMIND ^[1,8], RR ^[29], SER ^[30], TEM ^[31] and COPE ^[32] are all further replay-based methods. Among replay-based methods is the concern for storage of previous samples and the number of instances used in replay to cover the input space. Generative adversarial networks (GAN) have been used in generating samples, which could be used in conjunction with real images, and this can be used for replay, termed as *pseudo-rehearsal*, which has also been used in incremental learning applications ^{[33][34][35][36]}.

2.4.2. Regularisation Methods

Among the research in regularisation-based methods, De Lange et al. [17] divided it into two types: data-focused methods and prior-focused methods. Parameter regularisation and activation regularisation methods are the ones already researched to mitigate the problem of loss of previously learned information [37][38][39][40][41]. Data-focused methods use knowledge distillation from a previous model into the new model trained for the new task. The knowledge distillation process is used here to transfer knowledge, and it also helps mitigate the catastrophic forgetting during training for new tasks. The output from the model trained for a previous task is used as a soft label for those previously learned tasks. Regularisation-based methods vary the plasticity of weight depending on the importance of the new task as compared to the old task, previously learned [18]. Methods such as EWC [16], GEM ^[42] and A-GEM ^[43] use regularisation methods, but in GEM and A-GEM, regularisation along with replay methods are used. In EWC, weights are consolidated into a Fisher information matrix, which gives parameter uncertainty based on tractable approximation. Compared to EWC, synaptic intelligence (SI) [44] collects individual synapsesbased task-relevant information to produce a high-dimensional dynamic system of parameter space over the entire learning trajectory in the effort to alleviate catastrophic forgetting. MAS [45] is a regularisation-based method where changes to previously learned weights are regulated by penalising parameter changes using hyper-parameter optimisation, where importance of weights are estimated from the gradients of the squared L2-norm for output from previous tasks. IMM ^[46] matches moments of posterior distribution of neural networks in an incremental priority for first and second training tasks and uses further regularisation methods including weight transfer, L2-norm-based optimisation and also dropout methods ^[16].

2.4.3. Parameter Isolation Methods

Parameter isolation methods isolate and dedicate parameters for each specific task to mitigate forgetting. When new tasks are learned, previous tasks are masked either at unit level or at parameter level. This strategy of using a specific set of parameters for specific tasks may often lead to restricted learning capacity for new tasks ^[17]. New branches can be grown for new tasks given there is reduced constraint in network architecture. In such cases, parameter isolation is achieved by freezing previous task parameters ^{[47][48][49][50][51]}. PackNet ^[52] assigns network capacity to each task explicitly by using binary masks. In hard attention to task ^[20], task embedding is implemented by the addition of a fraction of previously learned weights to the network, which then computes conditioning mask using these high-attention vectors. The attention mask is used as a task identifier for each layer and utilises this information to prevent forgetting. In PNN ^[53], the weights are arranged in column-wise order respective to the new task with random initialisation of weights. Transfer between columns is enabled by adaptors, which are non-linear lateral connections between new columns created with each new task, thereby reducing catastrophic forgetting.

3. Incremental Learning for Edge Devices

In 2019, Li et al. studied incremental learning for object detection at the Edge ^[54]. The increased use of deep learning models for object detection on Edge computing devices was accelerated by one-stage detectors such as YOLO ^[55], SSD ^[56] and RetinaNet ^[57]. A deep learning model deployed at the Edge needs incremental learning to maintain the accuracy and robust performance in object detection in personalised applications. The algorithm termed as RILOD ^[54] is a method for incremental learning where the one stage detection network is trained end-to-end using a comparatively smaller number of images of the new class within the time span of a few minutes. The RILOD algorithm uses knowledge distillation, which has been used by several researchers to avoid catastrophic forgetting. Three types of knowledge from old models were distilled to mimic the output of previously learned classes on tasks such as object classification, bounding box regression and feature extraction.

In DeeSIL ^[58], fixed representations for class are extracted from a deep model and then used to learn shallow classifiers incrementally, which makes it an incremental learning adaptation for transfer learning. Since feature extractors replace real images, memory constrain for new data is addressed, hence making it a possible candidate for Edge devices. Train++ ^[59] is an incremental learning binary classifier for Edge devices, though it is based on training ML models on micro-controller units.

Rapid development of Edge Intelligence with optimisation of deep neural networks led to increased use of modelbased detection in computer vision applications. Due to its huge reduction in size as well as reduction in computational costs, the model quantisation is the most widely used optimisation method among all the other types of optimisation and compression techniques for deep learning. The detection models were trained on TensorFlow Object Detection API on a local GPU accelerated device running TensorFlow-GPU version 1.14, Python 3.6 and OpenCV 3.4 as the package for image analysis. The GPU used for training was GPU GeForce RTX 2080 Ti, after which TensorFlow-lite graph was exported as the frozen inference graph. This was then converted into flat-buffer format before integrating into Raspberry Pi 4, the resource constrained device used for detection experiments. The Raspberry Pi 4 runs Raspbian Buster-10 OS, with an integrated Raspberry Pi Camera Module V2 for real-time image capture in detection experiment. The Raspberry Pi camera is 8 megapixels, single channel, and has a maximum frame rate capture of 30 fps. The camera module connects to Raspberry Pi 4 via 15 cm ribbon cable which attaches the Pi Camera Serial Interface port (CSI) to the module slots on the Pi.

4. Process Prediction and Operator Training

Concept drift is another type of problem that occurs in continual learning ^{[15][60]}, which needs mention for the completeness of challenges in continual learning. Concept drift, also known as model drift, is the change in the statistical properties of a target variable that occurs due to the change in streaming data with the addition of new products/defects ^[61]. Data drift is found to be one of the factors leading to concept drift. Data drift is the change in distribution of input data instances, which result in variation of predictive results from the trained model. Concept drift can happen over time when the definition of an activity class previously learned might change in the future data streams when the newer models are trained from the streaming data. In the manufacturing setting, maintaining and improving machining efficiency is directly related to the quality of manufacturing end products ^[62].

Yu et al. studied process prediction in the aspect of milling stability and the effect of damping caused by tool wear in the manufacturing setting. This is an application of incremental learning from the sequential stream of data available and heavily based on concept drift, one of the imminent challenges in incremental learning. The concept drift in this application is the stability domain change, which in turn makes change to the stability boundary. Taking into consideration the time frame in which the sequential data are available, the concept drift is identified as four types:

- sudden drift: introduction of a new concept in a short time
- gradual drift: introduction of new concept gradually over a period of time to replace the old one
- incremental drift: the change of an old concept to a new concept, incrementally over a period of time
- reoccurring drift: an old concept reoccurs after a certain period of time.

Correction-based incremental learning augments negative samples into the training set, which were previously classified as positive samples (false positives) to improve the decision boundary. This experiment is also research in incremental learning but without convolutional neural networks for object detection applications. Fully connected layers in neural network architecture are trained in stochastic gradient descent manner with fewer samples that

strategically improve the decision boundary for the required task. In the study conducted by Ramos et al. ^[63], incremental learning based on artificial neural networks are again used to predict industrial electricity consumption by a facility using real-time data and forecasting algorithms. Sequential training data are updated every midnight during the forecast process, where the forecast process is supported by periods split by 5 min intervals.

When considering the paradigm shift from Industry 4.0 to Industry 5.0 use case studies, the impact on people and organisation as well as the technological advances that are proposed must be considered. The main implementation challenges that have been reported in Industry 4.0 applications have been in relation to security, resilience, the ability to withstand disruptions and catastrophic events, operator training and efficient use of digital data from sensors. Industry 4.0 and 5.0 are both aimed at an important dimension of efficient use of energy and technology ^[64]. Humans create and manage the production systems, hence humans are the main drivers of activities and creators of infrastructure, but the processes in the production will be automated, and any human operator will only be considered as the human-in-the-loop to assist the automated systems. This role of the operator will include selecting the samples from the sequential data acquired by the sensors and labelling or pre-processing for incremental learning techniques. Industry 5.0 prioritises human-machine interaction as opposed to the introduction of robots and automated systems into the manufacturing process in Industry 4.0 ^[65].

References

- Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A survey on deep transfer learning. In Proceedings of the Artificial Neural Networks and Machine Learning—ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, 4–7 October 2018; Proceedings, Part III 27. pp. 270–279.
- Wang, J.; Wang, X.; Shang-Guan, Y.; Gupta, A. Wanderlust: Online continual object detection in the real world. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 11–17 October 2021; pp. 10829–10838.
- 3. Qi, Z.; Xu, Y.; Wang, L.; Song, Y. Online multiple instance boosting for object detection. Neurocomputing 2011, 74, 1769–1775.
- Grabner, H.; Bischof, H. On-line boosting and vision. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; Volume 1, pp. 260–267.
- 5. Sahoo, D.; Pham, Q.; Lu, J.; Hoi, S.C. Online deep learning: Learning deep neural networks on the fly. arXiv 2017, arXiv:1711.03705.
- Han, H.; Yang, R.; Li, S.; Hu, R.; Li, X. SSGD: A smartphone screen glass dataset for defect detection. In Proceedings of the ICASSP 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 4–10 June 2023; pp. 1–5.

- 7. Ren, M.; Chen, N.; Qiu, H. Human-machine Collaborative Decision-making: An Evolutionary Roadmap Based on Cognitive Intelligence. Int. J. Soc. Robot. 2023, 15, 1101–1114.
- 8. Nunes, D.S.; Zhang, P.; Silva, J.S. A survey on human-in-the-loop applications towards an internet of all. IEEE Commun. Surv. Tutor. 2015, 17, 944–965.
- 9. Wang, J.; Ma, Y.; Zhang, L.; Gao, R.X.; Wu, D. Deep learning for smart manufacturing: Methods and applications. J. Manuf. Syst. 2018, 48, 144–156.
- Ren, J.; Ren, R.; Green, M.; Huang, X. Defect detection from X-ray images using a three-stage deep learning algorithm. In Proceedings of the 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), Halifax, NS, Canada, 5–8 May 2019; pp. 1–4.
- Bhatt, P.M.; Malhan, R.K.; Rajendran, P.; Shah, B.C.; Thakar, S.; Yoon, Y.J.; Gupta, S.K. Imagebased surface defect detection using deep learning: A review. J. Comput. Inf. Sci. Eng. 2021, 21, 040801.
- Tan, M.; Pang, R.; Le, Q.V. Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13– 19 June 2020; pp. 10781–10790.
- Babcock, B.; Babu, S.; Datar, M.; Motwani, R.; Widom, J. Models and issues in data stream systems. In Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, Madison, WI, USA, 3–5 June 2002; pp. 1–16.
- Alrawashdeh, K.; Purdy, C. Toward an online anomaly intrusion detection system based on deep learning. In Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016; pp. 195–200.
- 15. Hasan, M.; Roy-Chowdhury, A.K. A continuous learning framework for activity recognition using deep hybrid feature models. IEEE Trans. Multimed. 2015, 17, 1909–1922.
- Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; Veness, J.; Desjardins, G.; Rusu, A.A.; Milan, K.; Quan, J.; Ramalho, T.; Grabska-Barwinska, A.; et al. Overcoming catastrophic forgetting in neural networks. Proc. Natl. Acad. Sci. USA 2017, 114, 3521–3526.
- De Lange, M.; Aljundi, R.; Masana, M.; Parisot, S.; Jia, X.; Leonardis, A.; Slabaugh, G.; Tuytelaars, T. A continual learning survey: Defying forgetting in classification tasks. IEEE Trans. Pattern Anal. Mach. Intell. 2021, 44, 3366–3385.
- Hayes, T.L.; Kafle, K.; Shrestha, R.; Acharya, M.; Kanan, C. Remind your neural network to prevent catastrophic forgetting. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 466–483.
- 19. Mittal, S.; Galesso, S.; Brox, T. Essentials for class incremental learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25

June 2021; pp. 3513-3522.

- Serra, J.; Suris, D.; Miron, M.; Karatzoglou, A. Overcoming catastrophic forgetting with hard attention to the task. In Proceedings of the International Conference on Machine Learning, PMLR, Stockholm, Sweden, 10–15 July 2018; pp. 4548–4557.
- 21. Ramasesh, V.V.; Dyer, E.; Raghu, M. Anatomy of catastrophic forgetting: Hidden representations and task semantics. arXiv 2020, arXiv:2007.07400.
- Shmelkov, K.; Schmid, C.; Alahari, K. Incremental learning of object detectors without catastrophic forgetting. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 3400–3409.
- McCloskey, M.; Cohen, N.J. Catastrophic interference in connectionist networks: The sequential learning problem. In Psychology of Learning and Motivation; Elsevier: San Diego, CA, USA, 1989; Volume 24, pp. 109–165.
- Ramasesh, V.V.; Lewkowycz, A.; Dyer, E. Effect of scale on catastrophic forgetting in neural networks. In Proceedings of the International Conference on Learning Representations, Virtual, 3–7 May 2021.
- 25. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language models are unsupervised multitask learners. OpenAI Blog 2019, 1, 9.
- 26. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv 2018, arXiv:1810.04805.
- Rebuffi, S.A.; Kolesnikov, A.; Sperl, G.; Lampert, C.H. icarl: Incremental classifier and representation learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2001–2010.
- Rannen, A.; Aljundi, R.; Blaschko, M.B.; Tuytelaars, T. Encoder based lifelong learning. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1320–1328.
- 29. Rolnick, D.; Ahuja, A.; Schwarz, J.; Lillicrap, T.; Wayne, G. Experience replay for continual learning. In Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA, 2019; Volume 32.
- 30. Isele, D.; Cosgun, A. Selective experience replay for lifelong learning. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LO, USA, 2–7 February 2018; Volume 32.
- 31. Chaudhry, A.; Rohrbach, M.; Elhoseiny, M.; Ajanthan, T.; Dokania, P.K.; Torr, P.H.; Ranzato, M. On tiny episodic memories in continual learning. arXiv 2019, arXiv:1902.10486.
- 32. De Lange, M.; Tuytelaars, T. Continual prototype evolution: Learning online from non-stationary data streams. In Proceedings of the IEEE/CVF international Conference on Computer Vision,

Montreal, BC, Canada, 11–17 October 2021; pp. 8250–8259.

- 33. Shin, H.; Lee, J.K.; Kim, J.; Kim, J. Continual learning with deep generative replay. In Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA, 2017; Volume 30.
- 34. Atkinson, C.; McCane, B.; Szymanski, L.; Robins, A. Pseudo-recursal: Solving the catastrophic forgetting problem in deep neural networks. arXiv 2018, arXiv:1802.03875.
- 35. Lavda, F.; Ramapuram, J.; Gregorova, M.; Kalousis, A. Continual classification learning using generative models. arXiv 2018, arXiv:1810.10612.
- 36. Ramapuram, J.; Gregorova, M.; Kalousis, A. Lifelong generative modeling. Neurocomputing 2020, 404, 381–400.
- 37. Li, Z.; Hoiem, D. Learning without forgetting. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 40, 2935–2947.
- Liu, X.; Masana, M.; Herranz, L.; Van de Weijer, J.; Lopez, A.M.; Bagdanov, A.D. Rotate your networks: Better weight consolidation and less catastrophic forgetting. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 2262–2268.
- Chaudhry, A.; Dokania, P.K.; Ajanthan, T.; Torr, P.H. Riemannian walk for incremental learning: Understanding forgetting and intransigence. In Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 23–28 August 2018; pp. 532–547.
- 40. Jung, H.; Ju, J.; Jung, M.; Kim, J. Less-forgetting learning in deep neural networks. arXiv 2016, arXiv:1607.00122.
- Zhang, J.; Zhang, J.; Ghosh, S.; Li, D.; Tasci, S.; Heck, L.; Zhang, H.; Kuo, C.C.J. Classincremental learning via deep model consolidation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass Village, CO, USA, 1–5 March 2020; pp. 1131–1140.
- 42. Lopez-Paz, D.; Ranzato, M. Gradient episodic memory for continual learning. In Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA, 2017; Volume 30.
- 43. Chaudhry, A.; Ranzato, M.; Rohrbach, M.; Elhoseiny, M. Efficient lifelong learning with a-gem. arXiv 2018, arXiv:1812.00420.
- Zenke, F.; Poole, B.; Ganguli, S. Continual learning through synaptic intelligence. In Proceedings of the International Conference on Machine Learning, PMLR, Sydney, NSW, Australia, 6–11 August 2017; pp. 3987–3995.
- Aljundi, R.; Babiloni, F.; Elhoseiny, M.; Rohrbach, M.; Tuytelaars, T. Memory aware synapses: Learning what (not) to forget. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 139–154.

- 46. Lee, S.W.; Kim, J.H.; Jun, J.; Ha, J.W.; Zhang, B.T. Overcoming catastrophic forgetting by incremental moment matching. In Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA, 2017; Volume 30.
- 47. Fernando, C.; Banarse, D.; Blundell, C.; Zwols, Y.; Ha, D.; Rusu, A.A.; Pritzel, A.; Wierstra, D. Pathnet: Evolution channels gradient descent in super neural networks. arXiv 2017, arXiv:1701.08734.
- Mallya, A.; Davis, D.; Lazebnik, S. Piggyback: Adapting a single network to multiple tasks by learning to mask weights. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 67–82.
- Aljundi, R.; Chakravarty, P.; Tuytelaars, T. Expert gate: Lifelong learning with a network of experts. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3366–3375.
- 50. Xu, J.; Zhu, Z. Reinforced continual learning. In Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA, 2018; Volume 31.
- 51. Rosenfeld, A.; Tsotsos, J.K. Incremental learning through deep adaptation. IEEE Trans. Pattern Anal. Mach. Intell. 2018, 42, 651–663.
- 52. Mallya, A.; Lazebnik, S. Packnet: Adding multiple tasks to a single network by iterative pruning. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7765–7773.
- 53. Rusu, A.A.; Rabinowitz, N.C.; Desjardins, G.; Soyer, H.; Kirkpatrick, J.; Kavukcuoglu, K.; Pascanu, R.; Hadsell, R. Progressive neural networks. arXiv 2016, arXiv:1606.04671.
- 54. Li, D.; Tasci, S.; Ghosh, S.; Zhu, J.; Zhang, J.; Heck, L. RILOD: Near real-time incremental learning for object detection at the edge. In Proceedings of the 4th ACM/IEEE Symposium on Edge Computing, Arlington, VA, USA, 7–9 November 2019; pp. 113–126.
- 55. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
- Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Proceedings, Part I 14. pp. 21– 37.
- 57. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.

- Belouadah, E.; Popescu, A. DeeSIL: Deep-Shallow Incremental Learning. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018.
- Sudharsan, B.; Yadav, P.; Breslin, J.G.; Ali, M.I. Train++: An incremental ml model training algorithm to create self-learning iot devices. In Proceedings of the 2021 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI), Atlanta, GA, USA, 18–21 October 2021; pp. 97–106.
- 60. He, J.; Mao, R.; Shao, Z.; Zhu, F. Incremental learning in online scenario. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 13926–13935.
- 61. Lu, J.; Liu, A.; Dong, F.; Gu, F.; Gama, J.; Zhang, G. Learning under concept drift: A review. IEEE Trans. Knowl. Data Eng. 2018, 31, 2346–2363.
- 62. Yu, Y.Y.; Zhang, D.; Zhang, X.M.; Peng, X.B.; Ding, H. Online stability boundary drifting prediction in milling process: An incremental learning approach. Mech. Syst. Signal Process. 2022, 173, 109062.
- 63. Ramos, D.; Faria, P.; Vale, Z.; Mourinho, J.; Correia, R. Industrial facility electricity consumption forecast using artificial neural networks and incremental learning. Energies 2020, 13, 4774.
- 64. Zizic, M.C.; Mladineo, M.; Gjeldum, N.; Celent, L. From industry 4.0 towards industry 5.0: A review and analysis of paradigm shift for the people, organization and technology. Energies 2022, 15, 5221.
- 65. Alsamhi, S.H.; Shvetsov, A.V.; Kumar, S.; Hassan, J.; Alhartomi, M.A.; Shvetsova, S.V.; Sahal, R.; Hawbani, A. Computing in the sky: A survey on intelligent ubiquitous computing for uav-assisted 6g networks and industry 4.0/5.0. Drones 2022, 6, 177.

Retrieved from https://encyclopedia.pub/entry/history/show/124823