Editorial
Deep Neural Network Representation and Generative Adversarial Learning

Generative Adversarial Networks (GANs) have proven to be efficient systems for data generation and other machine learning tasks. They owe their success to a minimax learning concept initially proposed by Schmidhuber (1990) to implement Artificial Curiosity. Two learning networks, a generator and an evaluator or discriminator, compete with each other in a zero-sum game. Despite their obvious advantages and their application to a wide range of domains, GANs have yet to overcome several challenges such as non-convergence, overfitting, mode collapse, amongst others. New advancements in deep representation learning (RL) can help improve the learning process in Generative Adversarial Learning (GAL). For instance, RL can help address issues such as dataset bias and identify a set of features that are well suited for a given task.

This special issue on Deep RL and GAL provides a collection of high-quality articles addressing theoretical and practical aspects of deep RL and GAL. After a rigorous review process of 70 articles, 21 articles were selected for inclusion in this special issue. A brief summary of these selected articles is provided hereafter:

The popularity and outstanding performance of GANs is attributed to a minimax learning paradigm in which two networks compete witch each other in a zero-sum game. In "*Generative Adversarial Networks are special cases of Artificial Curiosity (1990) and also closely related to Predictability Minimization (1991)*", Schmidhuber looks into the origins of adversarial learning through a minimax learning approach and establishes that GANs are a special case of Artificial Curiosity. The author also points out that another method of 1991 called Predictability Minimization is based on a very similar minimax game.

The joint training of two models in GANs makes model interpretability a challenging task. In *Latent Dirichlet allocation based generative adversarial networks*, Pan et al. explore how to integrate GANs with data structure priors. The proposed methodology aims to shed some light into model interpretability by looking into how multiple generators correlate to data structure.

The difference in the dimensionality space between the source and target domains in GANs can lead to an unstable training process. In *Multi-projection of unequal dimension optimal transport theory for Generative Adversary Networks*, Lin et al. propose a novel approach in which the source and target probability measures are projected into low-dimensional subspaces. The authors treat this optimization problem as a multi-marginal optimal transport (OT) problem and provide explicit properties of the solution.

ORCID(s):

In addition to unstable training, the generator model in GANs is prone to model collapse, which results in failure to generate data with several variations. In *Analysis of the transferability and robustness of GANs evolved for Pareto set approximations*, Garciarena et al. propose a novel neuro-evolutionary approach to address mode collapse and demonstrate the scalability of their approach through a series of experiments. The authors propose a set of guidelines to find the best performing GAN architecture for a given task.

Image synthesis is one of the most popular applications of GANs. However, generating realistic images using GANs remains a challenge, particularly when specific features are required. In *Investigating object compositionality in Generative Adversarial Networks*, van Steenkiste et al. look at using object compositionality as an inductive bias for GANs. The authors propose using a number of generator models, with shared weights and separate latent vectors as input, to generate images as a composition of individual objects.

Generating multi-domain images is also a challenging task. In *Unsupervised multi-domain multimodal image-to-image translation with explicit domain-constrained disentanglement*, Xia et al. propose a framework for multi-domain image translation through a single network. Xia et al. show that disentanglement of content and style can lead to undesirable results. The proposed frameworks allows for generation of diverse outputs with unpaired training data. Similarly, the authors of *MetalGAN: Multi-domain label-less image synthesis using cGANs and meta-learning*, propose a multi-domain image synthesis approach also using a single network. In this paper, Fontanini et al. propose combing a conditional GAN along with Meta-Learning for image generation and domain switch, demonstrating promising results. In *CariGAN: Caricature generation through weakly paired adversarial learning*, Li et al. employ facial landmarks to constrain the generated caricature facial images and generate caricature images with reasonable facial deformation. The authors introduce a loss function to encourage diversity in the generator.

In *Lower dimensional kernels for video discriminators*, Kahembwe et al. look into the discriminator models used in GANs for video generation. The authors show that unconstrained discriminator models are difficult to optimize. The authors continue by proposing a methodology for the design of GAN discriminators for video generation, which improves the accuracy of contemporary state-of-the-art approaches.

Class imbalance is another common problem that leads to poor generalization performance in deep NNs. The authors of *CEGAN: Classification Enhancement Generative Adversarial Networks for Unraveling Data Imbalance Prob-*

lems propose generating new minority class samples through a so-called classification enhancement generative adversarial network (CEGAN). Similarly, in *Deep learning for symbols detection and classification in engineering drawings*, Elyan et al. use GANs to generate new training samples for minority classes of engineering drawings data and show improved classification accuracy. Additionally, Sáez Trigueros et al. generate new training samples to boost face recognition accuracy performance in their article *Generating photo-realistic training data to improve face recognition accuracy*.

In *Modality independent adversarial network for generalized zero shot image classification*, Zhang et al. propose using GANs within a zero-shot learning context to classify images of unseen target classes. The GAN model learns discriminative and high level cross-modal data representations. The authors perform a series of ablation studies to show the outstanding performance of the proposed approach.

Aside from image synthesis, image manipulation is a common application of GANs. In *Image manipulation with natural language using Two-sided Attentive Conditional Generative Adversarial Network*, Zhu et al. introduce a novel approach to attributed editing through natural language descriptions. This method is shown to be robust to scale invariance and outperforms contemporary methods. In *ClsGAN: Selective Attribute Editing Model based on Classification Adversarial Network*, Liu et al. propose a classification adversarial network that produces high resolution images and high attribute transformation accuracy. The authors show state-of-the-art performance in image quality and transfer accuracy through a series of ablation studies.

It is well-established in the literature that model convergence can be a challenge for GANs. The authors of *FPGAN: Face de-identification method with generative adversarial networks for social robots* propose an end-to-end framework that is guaranteed to convergence to a given criterion. Lin et al. apply the proposed methodology to face de-identification, and outline a metric for this task.

GANs are increasingly becoming popular in application domains where data is commonly complex. The authors of *High tissue contrast image synthesis via multistage attention-GAN: Application to segmenting brain MR scans* investigate the use of GANS to improve image contrast. Hamghalam et al. argue that their multi-stage approach reduces the gap between the source and target domains. The authors show outstanding performance of the proposed method when applied to multimodal image segmentation. In a similar application, the authors of in *GP-GAN: Brain tumor growth prediction using stacked 3D generative adversarial networks from longitudinal MR Images* propose a data-driven 3D GAN for growth prediction of Glioma tumors using magnetic resonance images. The proposed method by Elazab et al. can also be used for image segmentation. Furthermore, in *AMD-GAN: Attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images*, Xie et al. propose a GAN methodology for the detection of fundus disease from laser ophthalmocopy images (SLO). In this method, the generator model is guided by an autoencoder and the discriminator model is exploited to classify SLO images.

The minimax learning paradigm employed by GANs, introduced in early works on Artificial Curiosity (1990) and Predictability Minimization (1991), continues to gain momentum in other domain applications outside of image synthesis. In *Energy-efficient and damage-recovery slithering gait design for a snake-like robot based on reinforcement learning and inverse reinforcement learning*, Bing et al. propose using the generator model as the agent that learns to generate actions that produce optimal rewards. The discriminator model is used to rate whether an action was produced by the generator model, or from expert data. The authors show promising results in terms of energy efficiency and damage recovery.

In a different application, Javanmardi et al. propose a GAN based approach to tracking in *Appearance variation adaptation tracker using adversarial network*. The authors propose using the discriminator model to differentiate between recent and earlier target regions, and the generator model to produce an adaptation mask. With this formulation, the model is able to align feature vectors of recent and earlier target regions, as well as distinguish between background and foreground.

The articles presented here propose a variety of methods addressing some of the most complex challenges in RL and GAL. They also highlight that using GANs often requires designing specific model architectures and loss functions tailored to the task being solved. This special issue aims to promote and encourage the development of theoretical methods in deep RL and GAL in an attempt to advance this popular field.

The guest editors extend their gratitude to the authors, Editors in Chief and reviewers for their contributions and support in the realization of this special issue.

ARIEL RUIZ-GARCIA, *Guest Editor*
SeeChange.ai
Manchester, United Kingdom
ariel.9arcia@gmail.com

JÜRGEN SCHMIDHUBER, *Guest Editor*
NNAISENSE
Swiss AI Lab IDSIA
USI and SUPSI
Lugano, Switzerland
juergen@idsia.ch

VASILE PALADE, *Guest Editor*
Centre for Data Science
Coventry University
Coventry, United Kingdom
vasile.palade@coventry.ac.uk

CLIVE CHEONG TOOK, *Guest Editor*
Department of Electronic Engineering
Royal Holloway (University of London)
Surrey, United Kingdom
Clive.CheongTook@rhul.ac.uk

DANILO MANDIC, *Guest Editor*
Department of Electrical and Electronic Engineering
Imperial College London
London, United Kingdom
d.mandic@imperial.ac.uk