



A survey on text generation using generative adversarial networks

Gustavo H. de Rosa*, João P. Papa

Department of Computing São Paulo State University Bauru, Brazil

ARTICLE INFO

Article history:

Received 19 January 2021

Revised 10 May 2021

Accepted 3 June 2021

Available online 9 June 2021

Keywords:

Text generation

Generative adversarial Networks

Machine learning

Language modeling

Natural language processing

ABSTRACT

This work presents a thorough review concerning recent studies and text generation advancements using Generative Adversarial Networks. The usage of adversarial learning for text generation is promising as it provides alternatives to generate the so-called “natural” language. Nevertheless, adversarial text generation is not a simple task as its foremost architecture, the Generative Adversarial Networks, were designed to cope with continuous information (image) instead of discrete data (text). Thus, most works are based on three possible options, i.e., Gumbel-Softmax differentiation, Reinforcement Learning, and modified training objectives. All alternatives are reviewed in this survey as they present the most recent approaches for generating text using adversarial-based techniques. The selected works were taken from renowned databases, such as Science Direct, IEEEExplore, Springer, Association for Computing Machinery, and arXiv, whereas each selected work has been critically analyzed and assessed to present its objective, methodology, and experimental results.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

Deep Learning has received the spotlight due to its capacity to solve human-like tasks through hierarchical learning. Nevertheless, such an area depends on proper representation learning to obtain high-quality performance and accomplish the intended task. For instance, Zhang et al. [1] presented a survey on an effective paradigm for representation learning denoted as Concept Factorization, which is widely used to correlate data points and concepts as linear combinations. Furthermore, Zhang et al. [2] introduced a novel approach to enhance optical character recognition with residual-based representation learning, achieving state-of-the-art results using dense Convolutional Neural Networks.

Such applications are also extensible to the Natural Language Processing (NLP) area, which also depends on precise representation learning to understand how humans communicate using voice- and text-based patterns. Essentially, its ultimate goal is to create algorithms capable of interpreting the human language and between machines and humans.

A common task NLP-based task denoted as Language Modeling aims to understand the general text structure by learning its grammatical, syntactic, and semantical rules. From a computational point of view, this task aims at predicting a $t + 1$ timestep based on previous t timesteps, where each timestep can be composed

of characters, words, or even sentences. Thus, a well-learned language model can reproduce the training text structure by predicting new data based on past information, e.g., text generation. For example, imagine that a fictitious language composed of three tokens follows this structure: <subject> <verb> <noun>. If a language model is capable of learning such structure, it is also capable of predicting a <verb> token based on input <subject> token, as well as predicting a <noun> token based on input <subject> and <verb> tokens.

One may realize that such a task is notably different from traditional Machine Learning ones as it employs a new dependency, i.e., the time itself. In other words, tokens (characters, words, sentences) are interpreted as sequences, where each piece represents a particular timestep. Additionally, previous sequences must be carried to the future during learning, as the language model predicts data based on past information [3]. Traditionally, researchers focus on solving such a problem using Recurrent Neural Networks (RNN) [4], which can re-use prior knowledge to learn future information. Even though RNNs seem to be the ideal structure to tackle text generation problems, they are still susceptible to adversarial manipulation [5]. In other words, such models are vulnerable to slight modifications in data that were not introduced in the learning process, resulting in mispredictions and inefficient performance. This vulnerability represents potential security threats and may compromise real-world applications, remaining as an open research field.

Goodfellow et al. [6] proposed the Generative Adversarial Networks (GAN) and achieved several hallmarks concerning adver-

* Corresponding author.

E-mail addresses: gustavo.rosa@unesp.br (G.H. de Rosa), joao.papa@unesp.br (J.P. Papa).

serial learning. Some recent works [7–9] revealed their capability of learning data distributions and plausibly generating artificial image-based data. Notwithstanding, such networks were designed to deal with continuous information instead of discrete data [10], i.e., they are not suitable when applied with sequences of texts. Thus, several works attempt to tackle such an issue by employing Gumbel-Softmax differentiation [11], Reinforcement Learning (RL) [12], or modified training objectives [13]. This paper proposes to survey and discuss the approaches mentioned above applied explicitly to text generation tasks. Additionally, it aims to provide a comprehensive source of text-based GAN advancements, where architectures are critically analyzed in terms of datasets, objectives, evaluation metrics, and experimental results.

The remainder of this paper is organized as follows. Section 2 presents the survey structure, related surveys, and how surveyed works were selected. Section 3 presents some theoretical background concerning Natural Language Processing, Language Modeling, Generative Adversarial Networks, as well as some text generation datasets and topmost evaluation metrics. Section 4 discusses and provides an in-depth analysis of the surveyed works. Finally, Section 5 states overall conclusions and future insights regarding the surveyed context.

2. Survey structure

2.1. Related surveys

Text generation, formerly known as natural language generation, has received a great deal of attention throughout the last years, mainly due to the advance of computational power and Deep Learning-based architectures. Gatt and Krahmer [14] presented a survey regarding the advancements of natural language generation. They categorized and discussed state-of-the-art data-driven architectures and presented the main drawbacks related to the synergy between natural language generation and other Artificial Intelligence areas. Nevertheless, their survey aims to provide an overview of the task instead of diving into the models' architecture and their advantages/disadvantages.

A recent survey presented by Lu et al. [15] reviews and critically analyzes adversarial-based text models' development. Even though their survey only focuses on comparing six state-of-the-art architectures and the Maximum Likelihood Estimation over two benchmarking datasets, the authors established a reference for those wanting to work with text-based adversarial learning. Furthermore, Zhang et al. [16] proposed a survey in the context of adversarial attacks on Deep Learning models in Natural Language Processing, where they targeted on summarizing and discussing the types of adversarial attacks instead of discussing the latest advancements regarding Generative Adversarial Networks applied to text-based generation.

Therefore, the presented survey aims to fulfill the literature's blank spots and provide a unique reference of adversarial-based models applied to text generation. Extensive research has been conducted across several renowned databases to identify the most promising text-based Generative Adversarial Networks. Additionally, the selected models have been critically analyzed and discussed in terms of objective, methodology, experimental results, and evaluation metrics. Finally, the present work intends to categorize the text-based adversarial models according to their differentiability strategy¹, e.g., Gumbel-Softmax differentiation, Reinforcement Learning, and modified training objectives.

¹ It is essential to remember that Generative Adversarial Networks were not designed to work with discrete data.

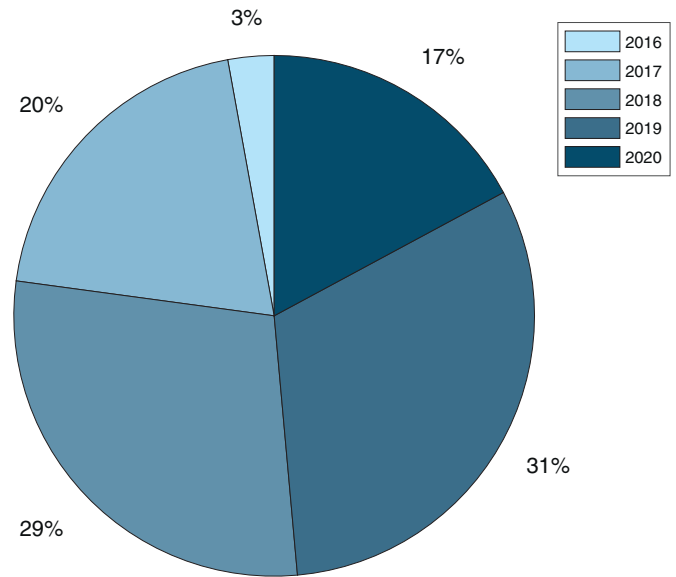


Fig. 1. Percentage of considered publications per publishing period.

2.2. Paper selection

Initially, we performed an extensive search at renowned databases, such as Science Direct, IEEEExplore, Springer, Association for Computing Machinery (ACM), and arXiv, to define the starting year of the surveyed works. As the Generative Adversarial Networks had been proposed in 2014 and the Gumbel-Softmax approach had been proposed in 2016, we could only find GAN-based text generation with Gumbel-Softmax references beyond 2016. Additionally, no work addressed GAN-based text generation with Reinforcement Learning and modified training objectives before 2017.

With that in mind, we considered two keywords for performing the search: (i) “generative adversarial network”, and (ii) “text generation”, and gathered the resulting works that have been cited at least one time and that were relevant to this survey context (GAN-based text generation). Furthermore, we verified their authors' background to check whether they were active researchers from truthful laboratories and universities. Thus, the following publication period is covered by the current survey: 2016 (1 work), 2017 (7 works), 2018 (10 works), 2019 (11 works), and 2020 (6 works), as illustrated by Fig. 1.

3. Theoretical background

This section presents a brief theoretical background regarding Natural Language Processing, Language Modeling, Generative Adversarial Networks, Reinforcement Learning, and Gumbel-Softmax differentiation-based strategies. Additionally, it describes an overview concerning the four most commonly employed text generation datasets and some topmost evaluation metrics.

3.1. Natural language processing

Most human communication is performed through voice and text, i.e., vital sources of information used to predict behavioral characteristics. Nevertheless, it is hard to predict human-based communication because each individual possesses particular traits in its written and verbal language, e.g., pitch, word selection, vocabulary, and accent.

An area that arises in tackling such a problem is Natural Language Processing, responsible for modeling and comprehending the interaction between humans and machines through natural lan-

guage usage. Deep Learning has recently achieved several hallmarks regarding NLP-based tasks [17], introducing state-of-the-art models for a wide variety of tasks, such as language modeling, grammatical analysis, and named entity recognition, among others.

Nonetheless, one may perceive that NLP is often complicated due to language's intrinsic features (morphology, syntax, semantic) and the requirement of knowledge to connect sequences of characters and words. Even though humans quickly learn and master languages, machines still struggle in such a task, leaving it as an open research problem for the community.

3.1.1. Language modeling

Language Modeling tasks stand for learning probabilistic models that can predict a $t + 1$ timestep given t timesteps, e.g., predict the next word given a sequence of words. A language model learns the probability of token occurrences based on examples, where simpler models might look at sequences of characters or words, while larger models may learn sequences of sentences or paragraphs. Notwithstanding, such tasks are often complicated as models need to learn natural languages' underlying traits, such as grammar, syntax, and semantic. Additionally, as its learning procedure is based on a set of examples, it poorly performs when employed with unknown data, e.g., not-learned tokens.

Most language models are used as root architectures for larger NLP tasks, such as speech recognition, optical character recognition, machine translation, and spelling correction. On the other hand, a trained language model can generate new text according to the structure it has already learned. Recently, most language models are fostered by neural networks, achieving state-of-the-art results due to their capacity for generalization. Furthermore, it is common to use pre-trained embeddings, i.e., real-valued vectors that encode high-dimensional representations of tokens (characters or words) in a projected vector space, to tackle the curse of dimensionality.²

One may perceive that an interesting approach is to use RNNs, as they can re-use their previous hidden states, where neurons are fed with recurrent connections and learn a so-called "memory". For example, a feed-forward network is not fit to create a probabilistic distribution capable of knowing that a <noun> should come after a <verb>, as it only looks into the "future". On the other hand, recurrent networks can look into their past "memory" and decide their subsequent output based on previous inputs.

3.1.2. How to generate text?

A properly trained language model encodes features and rules from the text it was trained. In other words, the model has learned a probabilistic distribution that represents the training data and samples new data. Let $P(w_t | w_{t-1}, w_{t-2}, \dots, w_{t-n})$ be the probability distribution of a token at t timestep, given an n -sized sequence of previous tokens. The language model estimates the probabilities of a token w_t given a previous sequence of tokens, allowing it to be sampled into a new token. Afterward, in $t + 1$ timestep, the probability distribution tries to represent w_{t+1} given $w_t, w_{t-1}, \dots, w_{t-n+1}$. The process is repeated until a convergence criterion is satisfied, such as the number of generated tokens. Finally, the resulting tokens are sequentially arranged to form the generated text.

3.2. Generative adversarial networks

Szegedy et al. [18] found in their work that several Machine Learning architectures, including state-of-the-art, are vulnerable to

adversarial manipulation, i.e., these models could not correctly classify when fed with slightly different learning data. Years ago, Goodfellow et al. [6] introduced the Generative Adversarial Networks as an alternative to the adversarial problem. It is implemented as a two-neural networks system, where the networks compete in a zero-sum approach. Namely, the idea is to have a discriminative and a generative network, where the generative one is in charge of producing fake data, and the discriminative one is in charge of estimating the probability of the fake data being real. Figure 2 illustrates an example of a standard Generative Adversarial Network.

The network's discriminator D is trained to maximize the probability of classifying both training and generated data as real images. Simultaneously, the network's generator G is trained to minimize $\log(1 - D(G(\mathbf{z})))$, i.e., the divergence among both data distributions, where \mathbf{z} denotes a noisy input. In other words, the two neural-networks system compete in a zero-sum game, represented by Eq. (1):

$$\min_G \max_D C(D, G) = \mathbb{E}_{\mathbf{x}}[\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z}}[\log(1 - D(G(\mathbf{z})))], \quad (1)$$

where $C(D, G)$ is the loss function to be minimized, $D(\mathbf{x})$ is the discriminator's estimated probability of a real sample \mathbf{x} being real, $\mathbb{E}_{\mathbf{x}}$ is the mathematical expectancy over all samples from the real data set \mathcal{X} , $G(\mathbf{z})$ stands for the generated fake data given the noise vector \mathbf{z} , $D(G(\mathbf{z}))$ is the estimated probability of a fake sample $G(\mathbf{z})$ being real, and $\mathbb{E}_{\mathbf{z}}$ is the mathematical expectancy over all random generator inputs, i.e., the expected value over all fake samples generated by G .

Nevertheless, a deriving problem from Eq. (1) is that GANs can get trapped in local optimums when the discriminator has an easy task. In other words, at the beginning of the training, when G still does not generate proper samples, D can reject fake samples with high probability, as they are incredibly different from the real data.³ Thus, an alternative to such a problem is to train G to maximize $\log(D(G(\mathbf{z})))$, which improves the initial gradient values.

3.2.1. Differentiating GANs with Gumbel-Softmax

During GANs training, both generator and discriminator need a differentiable loss function to update their parameters. However, when using RNN-based architectures as GANs generators, it is impossible to differentiate its output, usually accomplished by sampling one-hot categorical values from a Softmax distribution.

A solution to such a problem proposed by Jang et al. [11] consists of sampling from the so-called Gumbel-Softmax, which is a continuous distribution that can approximate samples from categorical distributions. Essentially, the idea is to employ the Softmax distribution along with a Gumbel distribution and a temperature parameter, as follows:

$$\tilde{\mathbf{y}} = \frac{e^{\frac{\mathbf{w} + \mathbf{g}}{\tau}}}{\sum e^{\frac{\mathbf{w} + \mathbf{g}}{\tau}}}, \quad (2)$$

where $\tilde{\mathbf{y}}$ stands for the differentiable outputs, \mathbf{w} is a vector of unnormalized log probabilities, \mathbf{g} is a vector sampled out of the Gumbel distribution, and τ is the temperature parameter. As $\tau \rightarrow 0$, $\tilde{\mathbf{y}}$ approximates a one-hot categorical distribution, while $\tilde{\mathbf{y}}$ approximates a uniform distribution as $\tau \rightarrow \infty$.

3.2.2. Reinforcing GANs learning

As previously mentioned, GANs need a differentiable loss function in order to perform their training. An alternative to the Gumbel-Softmax trick consists of creating a Reinforcement Learning-based agent for the generator. Essentially, an RL algorithm is defined by a policy function $\pi(\mathbf{a}|\mathbf{s}, \theta)$ parameterized by

² Curse of dimensionality happens when vast vocabularies are used, producing sparse representations.

³ Such a problem saturates the function $\log(1 - D(G(\mathbf{z})))$.

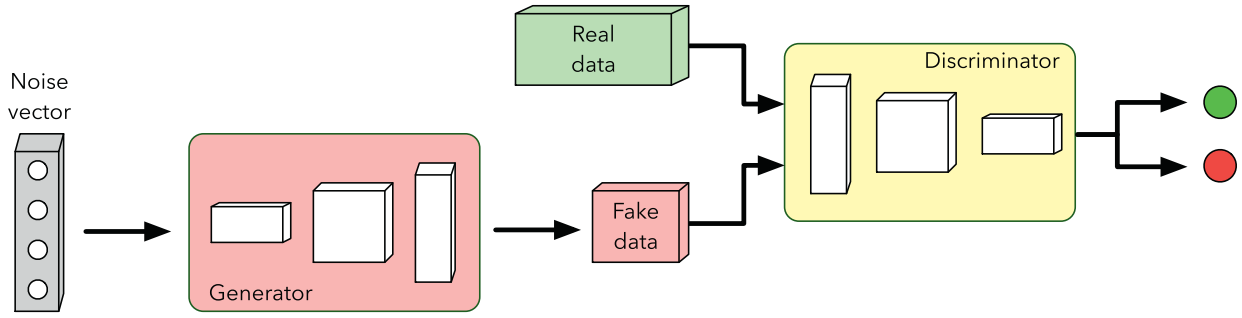


Fig. 2. Standard architecture of a Generative Adversarial Network.

θ , which outputs a probability distribution over all actions \mathbf{a} given the current state \mathbf{s} of an agent. Note that \mathbf{s} stands for the state of the agent (text generated), and \mathbf{a} stands for its actions (text to be generated). In other words, an agent generates text based on the policy π , where its actions are sampled according to the distribution defined by the same policy. Therefore, it is possible to optimize θ by employing policy gradient methods, such as the REINFORCE algorithm [19], to build a good text generation agent.

Let $\mathbf{y} = (y_1, y_2, \dots, y_T) \mid \mathbf{y} \in \mathcal{Y}$ depict the sequence generation problem, where \mathcal{Y} is the vocabulary of candidate tokens and T is the number of tokens to be generated. Let the generative model G be parametrized by θ_G and generate a state s for each time t , which corresponds to a generated token y_t . Further, let its next action, denoted by a , be the next token y_{t+1} to be chosen. Therefore, the policy $G(y_t \mid \mathbf{y}_{1:t-1}; \theta_G)$ is stochastic, such that the transition state is deterministic after an action has been chosen, i.e., $\delta_{s,s}^a = 1$ for the next state $\tilde{s} = \mathbf{y}_{1:t}$ if the current state $s = \mathbf{y}_{1:t-1}$ and the action $a = y_t$.

Additionally, the discriminative model D , which is parameterized by θ_D , guides the generator through a probability $D(\mathbf{y})$ that indicates whether a sequence \mathbf{y} belongs to real data or not. This discriminative model is trained by positive (real data) and negative (artificial data) samples, providing a reward signal to update the generative model's gradients. Eq. (3) describes the reward signal obtention procedure:

$$Q_D^G(a; s) = \begin{cases} \frac{1}{N} \sum_{n=1}^N D(\mathbf{y}^n) \mid \mathbf{y}^n \in MC(\mathbf{y}_{1:t}; N) & \text{for } t < T \\ D(\mathbf{y}) & \text{for } t = T \end{cases} \quad (3)$$

where $Q_D^G(a; s)$ is the action-value function of a sequence, i.e., the reward signal accumulated from state s given an action a , and N is the number of Monte Carlo samplings.

3.3. Datasets

The Amazon Customer Reviews⁴ [20] dataset, commonly denoted as Product Reviews, is composed of more than 233,000,00 users reviews, which have expressed their opinion and described their experience amongst products bought over Amazon's website. Furthermore, the Chinese Poems⁵ [21] dataset is composed of 284,299 poems from several online sources, such as Tang Poems, Song Poems, Song Ci, Ming Poems, Qing Poems, and Tai Poems.

The COCO Image Captions⁶ [22] dataset has been developed through the capture of image captions belonging to the Microsoft Common Objects in Context (MS COCO) dataset. In sum, humans created 1,026,459 captions from the same data sets used by MS COCO and divided them across two sub-sets: one with 5 reference sentences for each image and another one with 40 refer-

ence sentences for 5,000 randomly chosen images. All candidate captions have been pre-processed and tokenized according to the Stanford PTBTokenizer, following the same protocols established by the Penn Treebank dataset. After pre-processing, all punctuation marks have been removed to standardize the captions. Finally, the EMNLP2017 WMT News⁷ [23] dataset has been developed to foster machine translation tasks, being composed of pairs of documents, such as sets of sentences amongst two distinct languages. Nevertheless, it is possible to adjust such a dataset to the natural language generation task and use only sentences corresponding to the same language.

3.4. Evaluation metrics

The task of natural language generation allows the machine to create artificial information and understand natural languages. However, it is necessary to assess such information's quality and compare them against those created by human beings. Therefore, the advent of algorithms focused on evaluation metrics allowed the estimation of how "relevant" the generated information is and enabled ways to benchmark their quality between several language generation architectures.

3.4.1. Bilingual evaluation understudy

Bilingual Evaluation Understudy (BLEU) [24] evaluates the quality of machine-generated text and is structured by the correspondence of source text (written by humans) and artificial text (generated by machines). Essentially, it is calculated by the co-occurrence of n -grams between the reference and candidate sentences, computing a type of precision between the sentences. Let $\Phi = \{\phi_1, \phi_2, \dots, \phi_\kappa\}$ be the reference sentences with κ tokens, while $\tilde{\Phi} = \{\tilde{\phi}_1, \tilde{\phi}_2, \dots, \tilde{\phi}_\omega\}$ is the candidate sequence with ω tokens. Eq. (4) describes its formula, as follows:

$$P_n = \frac{\sum_k \min(d_k(\tilde{\phi}), \max d_k(\phi))}{\sum_k d_k(\tilde{\phi})}, \quad (4)$$

where $d_k(\cdot)$ is the amount of occurrences of an n -gram and k is the maximum number of possible n -grams. However, such formula inflates the metric's value for short sentences and needs a penalization factor according to Eq. (5):

$$b = \begin{cases} 1 & \text{if } \omega > \kappa \\ e^{1 - \frac{\kappa}{\omega}} & \text{if } \omega \leq \kappa \end{cases} \quad (5)$$

Therefore, the final BLEU score is calculated by a geometric mean weighted by the individual's n -grams precision. Moreover, the metric's value is within the $[0, 1]$ interval, where values close

⁴ <https://nijianmo.github.io/amazon/index.html>

⁵ <https://homepages.inf.ed.ac.uk/mlap/Data/EMNLP14>

⁶ <https://cocodataset.org>

⁷ <http://statmt.org/wmt17/results.html>

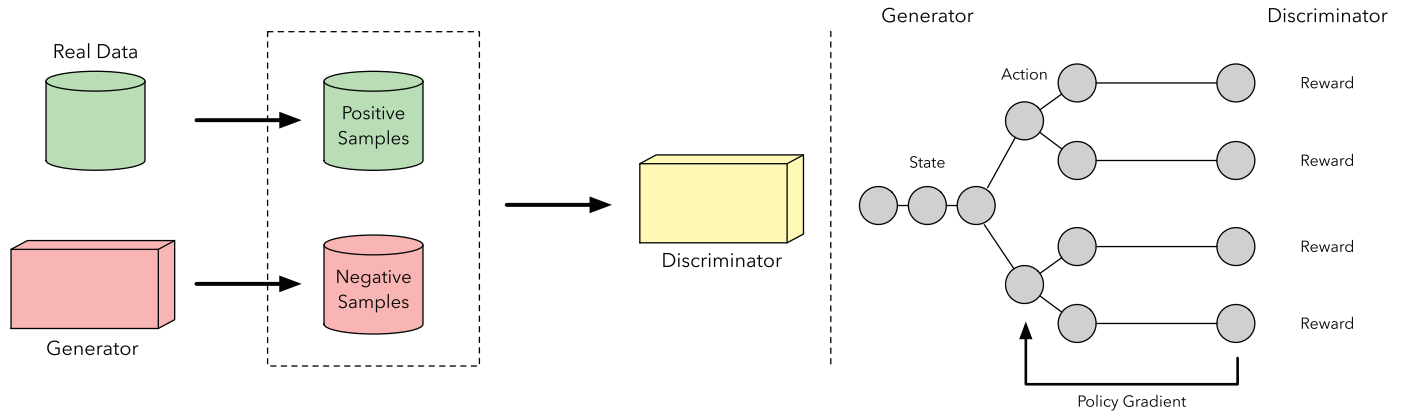


Fig. 3. Pipeline commonly used to conduct the adversarial-based text generation. Note that for Gumbel-Softmax differentiation and modified training objectives, the discriminator directly rewards the generator (without Reinforcement Learning) as their functions are differentiable.

to 1 indicate a bigger similarity between the sentences. Eq. (6) describes such procedure, as follows:

$$B_N = b \cdot \exp\left(\left(\sum_{n=1}^N w_n \log(P_n)\right)\right), \quad (6)$$

where $n = 1, 2, 3, \dots, N$ and w_n is a constant for all n values.

3.4.2. Negative log-likelihood

Negative Log-Likelihood (NLL) [25] is usually used as a loss function in Machine Learning training algorithms and aims at maximizing the probability of correct classifications. In other words, during a linguistic model training, NLL is used to verify whether the network is learning to generate $t + 1$ tokens given the previous t tokens.

Let K be the number of classes to be classified, i.e., the number of possible generated tokens. The loss function is defined according to Eq. (7):

$$NLL = -\frac{1}{n} \sum_{\mathbf{x} \in \mathcal{X}} \sum_{k=1}^K y_k \cdot e^{a_k}, \quad (7)$$

where n is the number of samples in training set \mathcal{X} , \mathbf{x} is the input data, \mathbf{y} is the true label in one-hot encoding, and a is the network's output. The NLL metric is commonly used to evaluate a network's learning quality and is restricted to a positive interval, e.g., $[0, +\infty]$. Additionally, as NLL tends to 0, the network can learn the training set patterns, i.e., correctly predicting the training samples' labels.

3.4.3. Perplexity

Perplexity (PPL) [26] is calculated by the exponentiation of NLL, furnishing a more intuitive metric. Given that PPL is just the exponentiation of NLL, its interval is restricted to $[1, +\infty]$, where smaller values correspond to a well-trained network, and higher values correspond to a poor learning procedure.

4. Analysis and discussion

This section presents the critical analysis and discussion regarding the compiled works, which are categorized as follows: (i) GAN-based text generation with Gumbel-Softmax differentiation, (ii) GAN-based text generation with Reinforcement Learning, and (iii) GAN-based text generation with modified training objectives. Additionally, Fig. 3 describes the methodological pipeline used to conduct text generation with the mentioned approaches, and Table 1 describes an overview regarding the surveyed works, where the following columns have depicted each work: reference, architecture, datasets, objectives, evaluation metrics, and experimental results, while Table 2 comprehends a set of architectures evaluated

with the most common metric (BLEU-2) over the four most common datasets: Amazon Reviews, Chinese Poems, COCO Image Captions, and EMNLP2017 WMT News.

4.1. Text generation using GANs and Gumbel-Softmax differentiation

Continuous distributions that resemble categorical distributions, such as the Gumbel-Softmax, tackled the differentiation problem and became alternatives to enable proper backpropagations of GANs' gradients. Initially, Kusner and Hernández-Lobato [27] presented the Gumbel-Softmax Generative Adversarial Network (GS-GAN), a text-based GAN that uses the Gumbel-Softmax distribution for generating its outputs. Their work was conducted using 5,000 samples with a maximum length of 12 characters generated from a context-free grammar (CFG) and evaluated by visual inspection of the generated data compared with a standard Maximum Likelihood Estimation (MLE) pre-trained LSTM. Nevertheless, such an architecture could only produce comparable results to the pre-trained LSTM and was no match for the Relational Generative Adversarial Network (RelGAN), introduced by Nie et al. [28].

The RelGAN introduced a Relational Memory-based generator, a Gumbel-Softmax relaxation, and a multi-embedded representation discriminator to compose their architecture. Their experiments were conducted using synthetic data generated by an oracle LSTM [30] and real datasets, such as COCO Image Captions and EMNLP2017 WMT News, being compared to state-of-the-art architectures (MLE, SeqGAN, RankGAN, and LeakGAN). Its main advantage lies in using a Relational Memory generator, which better extracts long-term dependencies due to self-attention layers. In contrast, its main disadvantage lies in needing a proper pre-training to generate feasible text, increasing the overall system's computational burden.

Nonetheless, adversarial-based models' increasing problem was the mode collapsing, where generators tend to sacrifice diversity for quality. In such a context, the Meta Cooperative Training Paradigm Generative Adversarial Network (Meta-CoTGAN), proposed by Yin et al. [29], attempted to slow the effect of mode collapsing. The authors evaluated the impact of the cooperative training language model and meta optimization using the same protocol proposed by Nie et al. [28]. Although the authors eased the mode collapsing problem, Meta-CoTGAN had a higher computational cost due to being trained with an additional language model during the adversarial training.

With the mentioned architecture in mind, it is possible to observe that their main disadvantage lies in pre-training both discriminator and generator ahead of the adversarial training, burdening the computational powers. On the other hand, one can perceive

Table 1
Summarization of the works considered in this survey.

Reference	Architecture	Dataset	Objectives	Metrics	Results
Kusner and Hernández-Lobato [27]	GSGAN	Context-Free Grammar	Gumbel-Softmax with discrete data.	Visual inspection	Samples similar to an MLE-trained LSTM.
Nie et al. [28]	RelGAN	COCO Image Captions, EMNLP2017 WMT News	Relational Memory generator and Gumbel-Softmax distribution.	NLL, BLEU	Outperformed models when RelGAN pre-training was good.
Yin et al. [29]	Meta-CoTGAN	COCO Image Captions, EMNLP2017 WMT News	Cooperative approach with an additional language model.	NLL, BLEU	Positive impact of cooperatively approach and meta optimization.
Yu et al. [30]	SeqGAN	Chinese Poems, Nottingham Music, Obama Speech, Synthetic Data	Stochastic Reinforcement Learning gradient policy updates.	NLL, BLEU	First model to extend GANs to discrete token generation.
Lin et al. [31]	RankGAN	Chinese Poems, COCO Image Captions, Shakespeare Plays	Ranked data instead of real and fake.	BLEU	Could not compete with human-written sentences.
Wang et al. [32]	VGAN	Taobao Reviews, Amazon Review, Penn Treebank	High-latent variables with Variational AutoEncoders (VAE).	NLL, BLEU	Successfully applied VAEs to adversarial-based models.
Guimaraes et al. [33]	ORGAN	Molecules Dataset, Essen Associative Code	Domain-specific objectives and discriminator's rewards.	Diversity, Validity	Generated domain-based data, with RL being essential.
Guo et al. [34]	LeakGAN	Chinese Poems, COCO Image Captions, EMNLP2017 WMT News, Synthetic Data	Leaked high-level features from the discriminator into the generator.	NLL, BLEU	Concluded that leaked information is vital to the network's learning.
Hjelm et al. [10]	BGAN	Google One Billion Words	Compute importance weights with difference measures.	Visual Inspection	Scratch stable training yet poor generation.
Fedus et al. [35]	MaskGAN	IMDB Movie, Penn Treebank	Actor-critic conditional architecture.	Percentage of Unique n -Grams	Produced reasonable sentences, yet not as good as the ground truth.
Li et al. [36]	CS-GAN	Amazon Review, Emotion, NEWS, Stanford Sentiment Treebank, Yelp Review	Incorporated sentence category over the architecture.	NLL, Accuracy	Suffered from large amounts of data and binary labels.
Balogopalan et al. [37]	RE-GAN	Context-Free Grammar	Comparative study between three distinct policy gradient methods.	Goodness Score	Needed a significant pre-training.
Shi et al. [38]	IRL	COCO Image Captions, IMDB Movie, Synthetic Data	Inverse RL architecture.	BLEU, F-BLEU, B-BLEU, HM-BLEU	Alleviated the problem of reward sparsity and mode collapse.
Xu et al. [39]	DP-GAN	Amazon Review, OpenSubtitles Dialogue, Yelp Review	Novel discriminator to distinguish between repeated and novel text.	BLEU, Relevance, Diversity, Fluency	Produced more diverse text, and better distinguished novel and repeated text.
Chen et al. [40]	CTGAN	Amazon Review, Film Review, Obama Speech, Yelp Review	Mimicked human-like text generation by introducing new features.	Similarity, Diversity, Word Frequency, Subjectivity, Sentiment	Simplified text generation process outperformed baselines.
Wang and Wan [41]	SentiGAN, C-SentiGAN	BeerAdvocate, Customer Reviews, Emotional Tweet, Stanford Sentiment Treebank	Multiple generators with a penalty-based objective.	Novelty, Diversity	Tagged as promising approaches in domain-based text generation.
Zhang et al. [42]	CD-GAN	Amazon Review, Yelp Review	Annotated data from distinct domains to generated emotional text.	Sentiment Transfer Strength, Domain Transfer Strength, Cosine Similarity, Word Overlap	Achieved comparable results to the compared baselines (sentiment transfer state-of-the-art).
Zhang et al. [43]	TranGAN	Penn Treebank	Transformer-based architecture trained with an actor-critic RL algorithm.	Perplexity, Percentage of Unique n -Grams, BLEU	Processed discrete sequences and outperformed compared baselines.
Liu et al. [44]	BFGAN	Amazon Review, Chinese Sentence Making Corpus, DailyDialog	Backward and forward generators with a discriminator capable of guiding their joint training.	BLEU, SLEU	Improved quality of lexically-constrained sequences, and did not need additional labels.
Sun et al. [45]	QuGAN	Tibetan Q&A	QRNNs and improved both reward and Monte Carlo search strategies.	BLEU	Model with optimized Monte Carlo search and BERT outperformed compared models.
Rizzo and Van [46]	C-SeqGAN	Amazon Review, COCO Image Captions, EMNLP2017 WMT News	Context adaption of global word embeddings, as well as a self-attention discriminator.	BLEU, Pos-BLEU, Self-BLEU	Improved SeqGAN by using global knowledge adapted to the dataset's domain.
Wu and Wang [47]	TG-SeqGAN	COCO Image Captions	Rewarded the generator with the distance between real and fake data.	NLL, Embedding Similarity	Improved the training convergence and the text quality.

(continued on next page)

Table 1 (continued)

Reference	Architecture	Dataset	Objectives	Metrics	Results
Zhou et al. [48]	SAL	COCO Image Captions, EMNLP2017 WMT News, Synthetic Data	Comparative discriminator instead of a binary one, as well as a self-improvement reward mechanism.	NLL, BLEU, Perplexity, Frechet Distance	Improved the rewards system by reducing the sparsity and mode collapse, leading to a better text generation.
Che et al. [13]	MaliGAN	Chinese Poems	Normalized maximum likelihood with importance sampling and variation reduction.	BLEU, Perplexity	Created distributions with lower variances and a better discriminator.
Zhang et al. [49]	TextGAN	Combined data from BookCorpus and ArXiv	High-dimensional latent feature distribution with kernelized discrepancy.	BLEU, Number of Kernel Density Estimators	Learned a suitable latent representation space and produced realistic sentences.
Gulrajani et al. [50]	WGAN-GP	Google One Billion Words	Gradient penalization.	Visual Inspection	Generated discrete data, yet not comparable to baselines.
Press et al. [51]	AltGAN	Google One Billion Words	Curriculum learning and slow teaching.	Percentage of Unique n -grams	Adequately trained from scratch.
Chen et al. [52]	FM-GAN	COCO Image Captions, CUB Captions, EMNLP2017 WMT News	Latent feature distributions with Feature-Mover's Distance.	Test-BLEU, Self-BLEU, Human Evaluation	Produced good sentences, yet still suffered from mode collapse.
Donahue and Rumshisky [53]	LaTextGAN	Toronto Book Corpus	AutoEncoder.	Human Discriminator, BLEU, t-SNE Plots	t-SNE plots showed that it could model the latent space.
Li et al. [54]	JSD-GAN	Chinese Poems, COCO Image Captions, EMNLP2017 WMT News, Obama Speech	Optimize Jensen-Shannon Divergence.	BLEU	RL-free with direct generator optimization, yet only for explicit representations.
Haidar et al. [55]	TextKD-GAN	Google One Billion Words, Stanford Natural Language Inference	Knowledge distillation technique based on teacher-student model.	BLEU	Smoother text representations and outperformed traditional GANs.
Ahamad [56]	STGAN	BookCorpus, CMU-SE	Encoder-decoder architecture inspired by the skip-gram model.	BLEU	Skip-thought embeddings could generate discrete data.
Montahaei et al. [57]	DGSAN	Chinese Poems, COCO Image Captions, EMNLP2017 WMT News	Optimized closed-form solution to find an optimal discriminator.	NLL, BLEU, Self-BLEU, MS-Jaccard	Outperformed baselines regarding the MS-Jaccard metric, yet achieved slightly lower Self-BLEU.
Yang et al. [58]	FGGAN	Chinese Poems, COCO Image Captions, Synthetic Data	Feature guidance module and a new vocabulary mask to better represent semantic rules.	NLL, BLEU	Outperformed current baselines and generated more realistic sentences due to token restriction.
Wu et al. [59]	TextGAIL	COCO Image Captions, EMNLP2017 WMT News	Pre-trained language models with contrastive discriminator and proximal policy optimization.	NLL, BLEU, Self-BLEU, Perplexity	Generated diverse and accurate outputs due to the incorporation of pre-trained GPT-2 and pre-trained RoBERTa.

Table 2

BLEU-2 benchmark regarding architectures that use such a metric to evaluate the four most common datasets.

Architecture	Amazon Review	Chinese Poems	COCO Captions	WMT News
MLE	-	0.667	0.781	0.768
RNNLM	0.848	-	-	-
RelGAN	-	-	0.849	0.881
Meta-CoTGAN	-	-	0.858	0.882
SeqGAN	0.856	0.739	0.745	0.777
VGAN	0.868	-	-	-
RankGAN	-	0.812	0.743	0.727
LeakGAN	-	0.881	0.746	0.826
IRL	-	-	0.829	-
BFGAN	0.920	-	-	-
SAL	-	-	0.785	0.788
FGGAN	-	-	0.773	-
MaliGAN	-	0.741	-	-
JSD-GAN	-	0.536	0.894	0.943

that Gumbel-Softmax distributions employ an additional temperature parameter, which controls the generated text's diversity and quality, thus being an interesting parameter to be fine-tuned. Additionally, such models mostly need a mathematical replacement in their output layer (from Softmax to Gumbel-Softmax), which provides a straightforward implementation and a not so complex system.

4.2. Text generation using GANs and reinforcement learning

Most Gumbel-Softmax-based approaches have a pre-training burden in advance to the adversarial training and directly rely on traditional GANs objectives, which may cause premature collapsing and an inadequate equilibrium between generator and discriminator. An alternative to tackling such a problem is to bypass the

generator's differentiation problem using stochastic gradient policy updates. The idea is to acquire rewards signals from the discriminator and pass them back to the generator in the same way a Reinforcement Learning method would do, such as the REINFORCE algorithm.

The first RL-based work has been introduced by Yu et al. [30], denoted as Sequence Generative Adversarial Network (SeqGAN). Their proposed architecture uses the REINFORCE algorithm, and Monte Carlo searches to propagate the discriminator's gradients to the generator. SeqGAN was the first RL-based work and could only be compared to standard MLE-based training, yet it established a hallmark for future RL-based research. Although SeqGAN tackled adversarial-based text generation, the model lacked generalization power and produced text with minimum diversity. With that in mind, an alternative suggested by Wang et al. [32] employs additional features learned by Variational AutoEncoders to represent text better. The so-called VGAN proposed to use high-level latent random variables to model the variability of text, aiding traditional Recurrent Neural Networks in learning well-structured data. Their combined model used an RL-based approach to update the generator's gradients and the same gradient policy defined by SeqGAN. Finally, VGAN outperformed the compared benchmarks, such as recurrent-based Language Models and SeqGAN, over three public literature datasets.

Another SeqGAN extension, denoted as Objective-Reinforced Generative Adversarial Networks (ORGAN), attempted to improve the learning process by including domain-specific objectives in addition to discriminator's rewards. In other words, ORGAN [33] proposed to extend the reward function through a linear combination between the discriminator and domain-specific objectives rewards, followed by a Wasserstein distance training. Their experimental results outperformed standard baseline approaches, such as MLE and SeqGAN, while the authors concluded that ORGAN could generate domain-based data and that RL plays an essential role in the model's learning.

Notwithstanding, such precedents enabled researchers to pursue new forms of inputting external information and improving domain-related text generation. In such a context, Li et al. [36] proposed the Category Sentence Generative Adversarial Network (CSGAN), which extends traditional text-based GANs with additional category information. Essentially, they proposed both generator and discriminator that incorporate the sentence category (label), allowing the system to capture such information in its learning procedure and generate text according to their category. On the other hand, Rizzo and Van [46] proposed an extension of the standard SeqGAN by employing context adaptation of global word embeddings and a self-attention discriminator. The so-called C-SeqGAN aimed at using the contribution of global knowledge adapted to the dataset's domain and a self-attentive discriminator that slowly minimizes the loss function and better provides feedback to the generator's updates. The experiments were performed under an extensive evaluation regarding short-, medium, and long-length text datasets and outperformed different SeqGAN-based baselines.

One can perceive the advantage of employing additional information, which assists the network in learning domain-related structures and producing more feasible texts. Nevertheless, using domain-related information may result in substantial mode collapses as the network will lose its diversity capability and generate similar text. Some authors opted not to rely on external information and employ variable reward signals to overcome such a problem. For instance, the Generative Adversarial Network (RankGAN) [31] employs a ranking-based architecture to produce high-quality generations. It evaluates and ranks collections of human- and machine-written sequences over a reference group instead of training a discriminator to learn binary predictions. One

can perceive that such an approach is inspired by the ranking steps commonly used in web searches and allows the model to calculate the expectations given different references sampled across the reference space. To calculate such values, the authors formulated a relevance score using cosine similarity followed by a softmax activation, which allowed them to compute a collective ranking score for an input sequence. Alternatively, the Diversity-Promoting Generative Adversarial Network (DP-GAN) [39] has been proposed to overcome the problem of generating repetitive and "boring" text sequences. Essentially, DP-GAN assigns a low reward value when generating text repeatedly while assigning high rewards when generating novel text. One can observe that such an approach encourages the generator to produce a more diverse text. Additionally, the authors proposed a new discriminator, which is better at distinguishing between repeated and novel text.

Moreover, the Boundary-Seeking Generative Adversarial Network (BGAN) [10] uses estimated difference measures from the discriminator to compute importance weights for the generated samples, providing a policy gradient for training the generator. In other words, BGAN models the strong connection between the discriminator's decision boundaries, allowing it to work with continuous data. BGAN uses f -divergences, such as Jensen-Shannon, Kullback-Leibler (KL), Reverse KL, or Squared-Hellinger policy gradients and rewards for the generator, being trained with the REINFORCE algorithm. BGAN yielded stable training even though its generation was inadequate as no pre-training was involved. The authors stated that they were not aware of any previous work that could successfully train from scratch a non-continuous architecture and that, although their generation was inadequate, the results demonstrated the stability and capacity of BGANs working with discrete data.

A recurrent disadvantage of modifying the rewards signals occurs when networks can not provide feasible signals during the starting epochs, hampering the learning process and preventing the generator from being adequately trained. In such a context, few works attempted to present novel RL-based strategies based on pre-training, joint rewards, and additional policy gradient algorithms. Liu et al. [44] presented a novel architecture for generating lexically-constrained sentences, denoted as Backward and Forward Generative Adversarial Network (BFGAN). Specifically, their idea was to employ both backward and forward generators with a discriminator that guides their joint training by assigning reward signals, being pre-trained with MLE before the adversarial learning, and then trained using the REINFORCE algorithm.

Regarding text-based GANs, policy gradient methods usually provide reward functions that sample a token per timestep without considering their surroundings and being unfeasible when dealing with long sequences. Thus, Balagopalan et al. [37] introduced a comparative study between three policy gradient alternatives, i.e., REBAR, RELAX, and slightly different REINFORCE. The proposed methods denoted as REGANs, aim at providing distinct reward functions to attain a probability distribution over all tokens generated by the generator. Experiments were conducted using synthetic context-free grammar and compared across REBAR, RELAX, and REINFORCE architectures. Alternatively, an Inverse Reinforcement Learning [60] algorithm has been proposed for text-based GANs, denoted as IRL [38]. The authors stated that their method has two main advantages over the traditional RL algorithm: (i) the reward function can produce denser reward signals, and (ii) an entropy regularized-based policy gradient encourages the generation of diverse texts. Unlike traditional GANs, IRL does not directly use a discriminator, where, instead, it uses a reward function that aims to increase the rewards from authentic texts and decrease the rewards from the generated texts. Thus, the generator is trained via entropy regularized policy gradients [61] by maximizing the expected rewards. The experimental results attest that the

proposed method alleviated reward sparsity and mode collapsing while achieving comparable results regarding the previous state-of-the-art, e.g., LeakGAN.

Most approaches mentioned above struggled to generate human-like texts and often faced a lack of diversity in the created sentences. With that in mind, some authors attempted to mimic human-like text generation by using previously unused features, such as variable text length, emotion label (positive, negative, or neutral), and controllable topic. The Conditional Text Generative Adversarial Network (CTGAN) [40] is trained using the REINFORCE algorithm and composed of a conditional LSTM generator that uses the emotion label and the text as its input. Additionally, it employed a conditional discriminator (standard CNN) to classify whether the text is real or generated. The authors customized the text generation using specific keywords and automated the word-level replacement strategy to extract the target word and replace it with a similar one.

Instead of only using emotions as additional features, Wang and Wan [41] extended the previous work by proposing both Sentimental Generative Adversarial Network (SentiGAN) and Conditional Sentimental Generative Adversarial Network (C-SentiGAN), which are capable of generating text with different sentiment labels. In short, the idea is to train multiple generators simultaneously and use a penalty-based objective function to force each generator in sampling specific sentiment texts. Furthermore, a multi-class discriminator helps each generator in generating its particular texts. Another architecture capable of generating emotional text is the Cross-Domain Text Sentiment Transfer Generative Adversarial Network (CD-GAN) [42], which uses annotated data from other domains and overcomes the problem of large-scale annotated data requirement. The CD-GAN aimed to combine adversarial RL and supervised learning to extract sentiment transformation patterns and generate emotional text. Its architecture comprises a Seq2Seq generator and two discriminators, one accountable for sentiments and another for domains. Both discriminators outputted rewards to update the generator's gradients and enabled the generator to observe how good its generation is based on the input data sentiment (non-labeled data) and related to the auxiliary domain (labeled data).

An alternative to traditional Reinforcement Learning methodologies is to split their knowledge into additional modules, such as manager/worker and actor/critic. Such architectures' advantage is their capacity to represent the policy functions better as they consider past actions and compare them to baselines to provide the reward signals. In other words, it decides how much reward should be given to a generator based on its previous text generations, guiding the network to more feasible updates and hence, better quality text. With that in mind, Guo et al. [34] proposed the Leaked Information Generative Adversarial Network (LeakGAN), which allows the discriminator to leak high-level extracted features to the generator and helps to guide the generator as it provides additional information through its learning steps. The idea is to split the generative network into two modules, the MANAGER and WORKER, where the MANAGER takes the leaked information from the discriminator and builds a feature vector further to feed the WORKER module. Both MANAGER and WORKER modules are distinct RNN-based architectures and trained via the REINFORCE algorithm. Additionally, the authors proposed a Bootstrapped Rescaled Activation, where rewards are re-scaled to mitigate the vanishing gradient problem. The experimental results showed that leaking information is vital to the network's learning process and better guides the generator in generating long text, achieving performance improvements over the previous state-of-the-art architectures.

Following the actor-critic-based architectures, Fedus et al. [35] introduced the Mask Generative Adversarial Net-

work (MaskGAN), which uses an actor-critic conditional GAN to fill in missing text conditioned on the surrounding context. One can perceive that such a task is more challenging than text generation, helping the model reduce the generator's risk competing with an almost-perfect discriminator. Additionally, in-filling tasks might mitigate the problem of mode collapsing as they provide rewards every timestep. The generator is based on a Seq2Seq [62] architecture and is trained via policy gradients using the REINFORCE algorithm. Furthermore, rewards are generated at each timestep, and the maximum sequence length is increased upon satisfying a convergence criterion. Also, to alleviate the REINFORCE variance problem, the authors proposed generating rewards based on the whole generator's distribution instead of only using the sampled token. On the other hand, Zhang et al. [43] improved the standard actor-critic methodology by proposing a transformer-based architecture for the generator, denoted as Transformer Generative Adversarial Network (TranGAN). Essentially, an encoder-decoder system constitutes the transformer-based architecture, where each layer is composed of multi-head self-attention and feed-forward networks in an attempt to extract long-term dependencies and information from the surrounding context. The proposed model, TranGAN, was compared against a standard Seq2Seq architecture and SeqGAN, outperforming both models in almost all metrics, while the authors concluded that they could effectively process discrete sequences through the actor-critic training algorithm.

Despite already known several RL-based strategies, researchers often grappled with the incapacity of generating long texts within a feasible amount of computation burden. Additionally, the lack of context (text semantics over time) in most GAN-based text generation architectures fostered the necessity of finding alternative improvements. Therefore, Sun et al. [45] presented a Quasi Generative Adversarial Network (QuGAN) for Tibetan Question Answering corpus generation, which extends traditional GANs by using the power of Quasi-Recurrent Neural Networks (QRNN) and improving both reward and Monte Carlo search strategies. As QRNNs are combinations of both RNNs and CNNs, they are often used to deal with long sequences of data and data parallelization, providing a faster text generation. Additionally, the authors employed a BERT model after the text generation to correct the generator's grammatical mistakes. Alternatively, by employing a truth-guided generator and self-attention-based discriminators, Wu and Wang [47] proposed the Truth-Guided Sequential Generative Adversarial Network (TG-SeqGAN) to improve text generation and make it look closer to the real data. The generator is essentially rewarded with the distance between real and fake data, guiding its convergence, while the discriminator network has a self-attention mechanism to capture long-term dependencies.

Another general improvement, proposed by the Self-Adversarial Learning (SAL) [48] architecture, is to employ a comparative discriminator capable of comparing text quality among samples' pairs. Moreover, throughout the training step, SAL rewards the generator when a generated sentence appears to be better than the previously generated ones. Such an approach stands for a self-improvement reward mechanism, which allows the model to receive more valuable feedback and avoid collapsing due to the limited number of real samples, thus reducing reward sparsity and the risk of mode collapse.

Finally, some new works attempt to improve the text generation quality by applying guidance functions to the discriminator's signals. Essentially, the idea is to provide more complex feedback and, consequently, a more feasible reward signal from the discriminator to the generator, leading to better and more diverse text. Yang et al. [58] introduced the Feature-Guiding Generative Adversarial Network (FGGAN), which solves the insufficient feedback guidance from the discriminator through a feature guidance mod-

ule. The module is a CNN that considers both generated and real data and performs a series of convolutional operations to extract a high-order feature vector and provide such a vector in the form of a reward signal. Additionally, the generated tokens pass through semantic rules to remove illogical tokens and improve the generated text's quality.

At the same time, the TextGAIL [59] attempts to improve the discriminator's guidance by combining large-scale pre-trained language models, i.e., RoBERTa and GPT-2, with recent advancements in Reinforcement Learning, such as Generative Information Learning (GAIL) and proximal policy optimization (PPO). The TextGAIL architecture extends the traditional adversarial learning with GAIL and introduces an imitation replay method to stabilize the training, i.e., vocabulary sizes are often large and hampers the GAIL stability. Additionally, ground truth sequences are filled in when training the generator, occasionally forcing constant rewards and stabilizing the generator's convergence. Their discriminator uses a pre-trained RoBERTa model as the final classifier and uses a contrastive-based approach that estimates the relative reality between generated and real sequences, i.e., the discriminator estimates how much a real sentence is more realistic than a generated sentence. As the discriminator is not stationary during adversarial learning, the generator suffers from high variance gradients and needs a type of policy optimization to stabilize its convergence, such as the PPO.

4.3. Text generation using GANs and modified training objectives

An alternative to continuous distributions and Reinforcement Learning is to propose new training objectives, hence withdrawing the necessity of differentiable Softmax-like, policy reward functions, and pre-training pipelines. The first work that attempted to introduce such a novel was Che et al. [13], which presented the Maximum-Likelihood Augmented Generative Adversarial Network (MaliGAN). Essentially, they introduced a novel objective based on the normalized maximum likelihood and importance sampling, which resembles an MLE-like objective and brings more stability to the training procedure. Additionally, one can observe that their variance reduction techniques assisted create discriminators capable of detecting whether a generator has learned too much noise.

A different architecture that proposes variance reduction techniques and aims at stabilizing the training procedure is the Improved Wasserstein Generative Adversarial Network (IWGAN) [50], which uses a gradient penalty technique instead of the standard weight clipping. Wasserstein networks commonly use weight clipping to stabilize their training, yet in a hard-optimization task, such a technique may not converge under a specific clipped value. In contrast, a gradient penalization is an alternative method to enforce the Lipschitz constraint by introducing a penalization factor to the critic's (discriminator) loss function. Unfortunately, there were not enough baselines to compare and attest whether IWGAN would scale with larger language models.

The literature also started to employ curriculum learning to mitigate the variance noise, where the algorithm slowly teaches the model how to generate increasing and variable-length sequences. Such a model denoted as AltGAN [51] uses a continuous relaxation mechanism, where the softmax outputs are multiplied with their respective embeddings to provide a differentiable generator. Their experimental results attested that their model could learn from scratch without resorting to any pre-training, being the first model to accomplish such a fact.

Another alternative approach to modify the training objective is by encoding sentences in latent feature distributions similar to AutoEncoders. The first work in such a category has been proposed by Zhang et al. [49], which introduced a text-based Generative Adversarial Network (TextGAN) that matches the high-dimensional latent

feature distribution of real and fake sentences using a kernelized discrepancy metric known as Reproducing Kernel Hilbert Space (RKHS). Such a model used a soft-argmax operator [63] as well as several complementary techniques, such as initialization strategies and discretization approximations, to alleviate the mode collapse problem. Donahue and Rumshisky [53] extended the previous work and proposed a latent-space text-based Generative Adversarial Network (LaTextGAN), which uses an AutoEncoder and low-dimensional latent spaces. In short, they attempt to remove the dimensionality curse and provide an encoding-decoding system that can generate low-dimensional latent-based spaces to discriminate the sequences better. Some plotted t-SNE embeddings showed that LaTextGAN could adequately model the latent space, while interpolated sentences indicated that LaTextGAN could smoothly encode sentences in the latent space.

Ahamad [56] introduced a Skip-Thought Vector Generative Adversarial Network (STGAN), which uses an encoder-decoder architecture to generate sentence embeddings. The proposed method does not resort to soft-differentiation or RL, as it generates text by feeding continuous embedded samples. Additionally, it did not have a specific training objective, which allowed the authors to compare and benchmark several training algorithms, such as mini-batch discrimination, gradient penalty, WGAN, and WGAN-GP. Following the STGAN proposal, the Text Knowledge Distillation Generative Adversarial Network (TextKD-GAN) [55] also aimed at producing continuous text representation to replace the traditional one-hot encodings. Instead of only producing embeddings with AutoEncoders, the TextKD-GAN resorted to a teacher-student model, feeding the data through an AutoEncoder and further reconstructing it. Such an approach provides a more challenging problem to the discriminator and increases the generator's ability to fool the discriminator.

Alternatively, the literature suggested using distance-based functions instead of encoder-decoder architectures to provide differentiable and straightforward objectives. In such a context, the Feature Mover Generative Adversarial Network (FM-GAN) [52] matches the latent feature distributions of real and fake sentences using Feature-Mover's Distance (FMD), leading to highly discriminative critics and RL-free approaches. The FMD is based on a variation of the Earth-Mover's Distance (EMD), which can be solved by the Inexact Proximal algorithm and further be applied to model discrete data objectives. This architecture could produce good sentences but still suffered from mode collapse (low diversity), while the compared baselines produced high diversity yet low-quality sentences. In such a context, Li et al. [54] presented an alternative distance-based architecture, the Jensen-Shannon Divergence Generative Adversarial Network (JSD-GAN). Such a model optimizes the Jensen-Shannon Divergence between the generator's distribution and the empirical distribution over the training data, i.e., an alternative to the min-max optimization, which uses a closed-form solution for the discriminator. Although their architecture provides an RL-free algorithm, they could only apply it to explicit representations, thus hindering the number of benefited tasks.

In contrast, some authors found out that using closed-form solutions directly optimizes the training objectives. For instance, Montahaei et al. [57] introduced the Discrete Generative Self-Adversarial Network (DGSAN), which uses an optimized closed-form solution to find an optimal discriminator. Essentially, the DGSAN combines both generator and discriminator powers in a single network, where the information is encoded into a new objective that creates a relationship between the current generator, the new generator, and the optimal discriminator between them. In short, the proposed method takes advantage of not having to calculate gradients, allowing for a direct sampling over an explicit distribution.

5. Conclusions and future insights

Throughout the last years, increasing demand for understanding and automating human communication fostered Natural Language Processing research. Machine Learning and abundant computational power helped achieve several hallmarks in a wide range of NLP-based tasks, thus becoming viable alternatives for modeling human-machine interactions.

Nevertheless, such systems still struggle when dealing with adversarial or unknown data, i.e., data that has been slightly modified or not present in the original one. On the other hand, one can perceive that such a problem encourages researchers to find feasible alternatives, such as Adversarial Learning. In essence, Adversarial Learning attempts to overcome the problem of adversarial data by introducing noisy examples to an architecture's learning process, aiding in its robustness and performance. Moreover, since the advent of Generative Adversarial Networks, it is now possible to learn the data's distribution and generate artificial data that resembles the original one.

This article presented a survey on the most recent studies concerning text generation using Generative Adversarial Networks. This paper's most significant contribution is to critically analyze and provide a unique source of recent GAN-based text generation research, mostly ranging from 2016 to 2020. Additionally, Table 1 presents an overview of the considered works, easing the reader's needs when looking for particular researches.

We have observed that GAN-based text generation works are brand new (between 2016 and 2020); hence they still struggle with a lack of researches and adequately defined pipelines. GAN-based text generation's primary challenges are that GANs were not designed to work with discrete data, needing additional tricks to subdue such a problem. Such data do not provide differentiable outputs, which inhibits gradients from being appropriately calculated and updated. Therefore, most works attempt to override such a problem by employing modified training objectives, Reinforcement Learning, or continuous-based outputs, such as Soft-Argmax or Gumbel-Softmax distributions.

Another barrier regards the intrinsic features of a language, i.e., grammar, syntax, and semantic properties. To provide a feasible text generation, the model needs to learn how characters and words connect between themselves, often accomplished by sorts of memories and contexts (prior knowledge). We believe that such problems can be dealt with in a more robust pre-learning step, in which pre-trained embedding models guide networks instead of MLE, such as BERT [64], ELECTRA [65], and GPT-2 [66], among others.

The advent of Transformer-based architectures, which provides substantial methods that can generate plausible "natural" language, initially diverted GAN-based systems' attention; however, their outstanding performance might benefit GANs when applied as pre-training approaches or even as part of their architectures. With that in mind, it seems that there are loads of blank spots that still need to be appropriately addressed by the research community regarding GAN-based text generation, whereas Transformer-based GANs look like promising approaches to be conducted in 2021.

Declaration of Competing Interest

None.

Acknowledgments

The authors are grateful to São Paulo Research Foundation (Fapesp) grants #2013/07375-0, #2014/12236-1, #2019/07665-4, #2019/02205-5, and #2020/12101-0, and to the Brazilian National Council for Research and Development (CNPq) #307066/2017-7

and #427968/2018-6. The authors are grateful to São Paulo Research Foundation (FAPESP) grant #2019/02205-5.

References

- [1] Z. Zhang, Y. Zhang, M. Xu, L. Zhang, Y. Yang, S. Yan, A survey on concept factorization: from shallow to deep representation learning, *Inf. Process. Manage.* 58 (3) (2021) 102534.
- [2] Z. Zhang, Z. Tang, Y. Wang, Z. Zhang, C. Zhan, Z. Zha, M. Wang, Dense residual network: enhancing global dense feature flow for character recognition, *Neural Netw.* 139 (2021) 77–85.
- [3] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, S. Khudanpur, Recurrent neural network based language model, in: 11th Annual Conference of the International Speech Communication Association, 2010.
- [4] J.L. Elman, Finding structure in time, *Cognit. Sci.* 14 (2) (1990) 179–211.
- [5] B. Biggio, F. Roli, Wild patterns: ten years after the rise of adversarial machine learning, *Pattern Recognit.* 84 (1) (2018) 317–331.
- [6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: *Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680.
- [7] X. Hu, P. Ma, Z. Mai, S. Peng, Z. Yang, L. Wang, Face hallucination from low quality images using definition-scalable inference, *Pattern Recognit.* 94 (1) (2019) 110–121.
- [8] D. Li, C. Du, H. He, Semi-supervised cross-modal image generation with generative adversarial networks, *Pattern Recognit.* 100 (1) (2020) 107085.
- [9] L. Liu, H. Zhang, X. Xu, Z. Zhang, S. Yan, Collocating clothes with generative adversarial networks cosupervised by categories and attributes: a multidiscriminator framework, *IEEE Trans. Neural Netw. Learn. Syst.* 31 (9) (2020) 3540–3554.
- [10] R.D. Hjelm, A.P. Jacob, T. Che, A. Trischler, K. Cho, Y. Bengio, Boundary-seeking generative adversarial networks, in: 6th International Conference on Learning Representations, 2018.
- [11] E. Jang, S. Gu, B. Poole, Categorical reparameterization with Gumbel-Softmax, in: 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017, Conference Track Proceedings, 2017.
- [12] R.S. Sutton, A.G. Barto, et al., *Introduction to Reinforcement Learning*, vol. 135, MIT Press Cambridge, 1998.
- [13] T. Che, Y. Li, R. Zhang, R.D. Hjelm, W. Li, Y. Song, Y. Bengio, Maximum-likelihood augmented discrete generative adversarial networks, 2017, *arXiv preprint arXiv:1702.07983*.
- [14] A. Gatt, E. Krahmer, Survey of the state of the art in natural language generation: core tasks, applications and evaluation, *J. Artif. Intell. Res.* 61 (2018) 65–170.
- [15] S. Lu, Y. Zhu, W. Zhang, J. Wang, Y. Yu, Neural text generation: past, present and beyond, 2018, *arXiv preprint arXiv:1803.07133*.
- [16] W. Zhang, Q. Sheng, A. Alhazmi, C. Li, Adversarial attacks on deep-learning models in natural language processing: a survey, *ACM Trans. Intell. Syst. Technol.* 11 (3) (2020) 1–41.
- [17] Y. Goldberg, A primer on neural network models for natural language processing, *J. Artif. Intell. Res.* 57 (1) (2016) 345–420.
- [18] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, R. Fergus, Intriguing properties of neural networks, in: *International Conference on Learning Representations*, 2014.
- [19] R.J. Williams, Simple statistical gradient-following algorithms for connectionist reinforcement learning, *Mach. Learn.* 8 (3–4) (1992) 229–256.
- [20] J. Ni, J. Li, J. McAuley, Justifying recommendations using distantly-labeled reviews and fine-grained aspects, in: 9th International Joint Conference on Natural Language Processing, 2019, pp. 188–197.
- [21] X. Zhang, M. Lapata, Chinese poetry generation with recurrent neural networks, in: *Conference on Empirical Methods in Natural Language Processing*, 2014, pp. 670–680.
- [22] X. Chen, H. Fang, T.-Y. Lin, R. Vedantam, S. Gupta, P. Dollar, C.L. Zitnick, Microsoft COCO captions: data collection and evaluation server, 2015, *arXiv preprint arXiv:1504.00325*.
- [23] B. Ondrej, R. Chatterjee, F. Christian, G. Yvette, H. Barry, H. Matthias, K. Philipp, L. Qun, L. Varvara, M. Christof, et al., Findings of the 2017 conference on machine translation (WMT17), in: 2nd Conference on Machine Translation, 2017, pp. 169–214.
- [24] K. Papineni, S. Roukos, T. Ward, W.-J. Zhu, BLEU: a method for automatic evaluation of machine translation, in: 40th Annual Meeting on Association for Computational Linguistics, 2002, pp. 311–318.
- [25] I. Myung, Tutorial on maximum likelihood estimation, *J. Math. Psychol.* 47 (1) (2003) 90–100.
- [26] P. Brown, S.D. Pietra, V.D. Pietra, J. Lai, R. Mercer, An estimate of an upper bound for the entropy of english, *Comput. Linguist.* 18 (1) (1992) 31–40.
- [27] M.J. Kusner, J.M. Hernández-Lobato, GANs for sequences of discrete elements with the Gumbel-Softmax distribution, 2016, *arXiv preprint arXiv:1611.04051*.
- [28] W. Nie, N. Narodytska, A. Patel, RelGAN: relational generative adversarial networks for text generation, in: 7th International Conference on Learning Representations, 2019.
- [29] H. Yin, D. Li, X. Li, P. Li, Meta-CoTGAN: a meta cooperative training paradigm for improving adversarial text generation, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, pp. 9466–9473.
- [30] L. Yu, W. Zhang, J. Wang, Y. Yu, SeqGAN: sequence generative adversarial nets with policy gradient, in: 31st AAAI Conference on Artificial Intelligence, 2017.

- [31] K. Lin, D. Li, X. He, Z. Zhang, M.-T. Sun, Adversarial ranking for language generation, in: 31st International Conference on Neural Information Processing Systems, 2017, pp. 3158–3168.
- [32] H. Wang, Z. Qin, T. Wan, Text generation based on generative adversarial nets with latent variables, in: *Advances in Knowledge Discovery and Data Mining*, Springer International Publishing, Cham, 2018, pp. 92–103.
- [33] G.L. Guimaraes, B. Sanchez-Lengeling, C. Outeiral, P.L.C. Farias, A. Aspuru-Guzik, Objective-reinforced generative adversarial networks (organ) for sequence generation models, 2018, arXiv preprint arXiv:1705.10843.
- [34] J. Guo, S. Lu, H. Cai, W. Zhang, Y. Yu, J. Wang, Long text generation via adversarial training with leaked information, in: 32nd AAAI Conference on Artificial Intelligence, 2018.
- [35] W. Fedus, I.J. Goodfellow, A.M. Dai, MaskGAN: better text generation via filling in the _____, in: 6th International Conference on Learning Representations, 2018.
- [36] Y. Li, Q. Pan, S. Wang, T. Yang, E. Cambria, A generative model for category text generation, *Inf. Sci.* 450 (1) (2018) 301–315.
- [37] A. Balagopalan, S. Gorti, M. Ravaut, R. Saqr, ReGAN: RE[LAX|BAR|INFORCE] based sequence generation using GANs, 2018, arXiv preprint arXiv:1805.02788.
- [38] Z. Shi, X. Chen, X. Qiu, X. Huang, Toward diverse text generation with inverse reinforcement learning, in: 27th International Joint Conference on Artificial Intelligence, 2018, pp. 4361–4367.
- [39] J. Xu, X. Ren, J. Lin, X. Sun, Diversity-promoting GAN: a cross-entropy based generative adversarial network for diversified text generation, in: *Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 3940–3949.
- [40] J. Chen, Y. Wu, C. Jia, H. Zheng, G. Huang, Customizable text generation via conditional text generative adversarial network, *Neurocomputing* 416 (1) (2020) 125–135.
- [41] K. Wang, X. Wan, Automatic generation of sentimental texts via mixture adversarial networks, *Artif. Intell.* 275 (1) (2019) 540–558.
- [42] R. Zhang, Z. Wang, K. Yin, Z. Huang, Emotional text generation based on cross-domain sentiment transfer, *IEEE Access* 7 (2019) 100081–100089.
- [43] C. Zhang, C. Xiong, L. Wang, A research on generative adversarial networks applied to text generation, in: 14th International Conference on Computer Science Education (ICCSE), 2019, pp. 913–917.
- [44] D. Liu, J. Fu, Q. Qu, J. Lv, BFGAN: backward and forward generative adversarial networks for lexically constrained sentence generation, *IEEE/ACM Trans. Audio Speech Lang. Process.* 27 (12) (2019) 2350–2361.
- [45] Y. Sun, C. Chen, T. Xia, X. Zhao, QuGAN: quasi generative adversarial network for tibetan question answering corpus generation, *IEEE Access* 7 (2019) 116247–116255.
- [46] G. Rizzo, T.H.M. Van, Adversarial text generation with context adapted global knowledge and a self-attentive discriminator, *Inf. Process. Manage.* 57 (6) (2020) 102217.
- [47] Y. Wu, J. Wang, Text generation service model based on truth-guided seqGAN, *IEEE Access* 8 (2020) 11880–11886.
- [48] W. Zhou, T. Ge, K. Xu, F. Wei, M. Zhou, Self-adversarial learning with comparative discrimination for text generation, 2020, arXiv preprint arXiv:2001.11691.
- [49] Y. Zhang, Z. Gan, K. Fan, Z. Chen, R. Henao, D. Shen, L. Carin, Adversarial feature matching for text generation, in: 34th International Conference on Machine Learning, 2017, pp. 4006–4015.
- [50] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A.C. Courville, Improved training of Wasserstein GANs, in: *Advances in Neural Information Processing Systems*, 2017, pp. 5767–5777.
- [51] O. Press, A. Bar, B. Bogin, J. Berant, L. Wolf, Language generation with recurrent generative adversarial networks without pre-training, 2017, arXiv preprint arXiv:1706.01399.
- [52] L. Chen, S. Dai, C. Tao, H. Zhang, Z. Gan, D. Shen, Y. Zhang, G. Wang, R. Zhang, L. Carin, Adversarial text generation via feature-mover's distance, in: *Advances in Neural Information Processing Systems*, 2018, pp. 4666–4677.
- [53] D. Donahue, A. Rumshisky, Adversarial text generation without reinforcement learning, 2018, arXiv preprint arXiv:1810.06640.
- [54] Z. Li, T. Xia, X. Lou, K. Xu, S. Wang, J. Xiao, Adversarial discrete sequence generation without explicit neural networks as discriminators, in: 22nd International Conference on Artificial Intelligence and Statistics, 2019, pp. 3089–3098.
- [55] M. Haidar, M. Rezagholizadeh, A.D. Omri, A. Rashid, Latent code and text-based generative adversarial networks for soft-text generation, in: 2019 Conference of the North American Chapter of the Association for Computational Linguistics, 2019, pp. 2248–2258.
- [56] A. Ahamad, Generating text through adversarial training using skip-thought vectors, in: Conference of the North American Chapter of the Association for Computational Linguistics, 2019, pp. 53–60.
- [57] E. Montahaei, D. Alihosseini, M. Soleymani Baghshah, DGSAN: discrete generative self-adversarial network, *Neurocomputing* 448 (2021) 364–379.
- [58] Y. Yang, X. Dan, X. Qiu, Z. Gao, FGGAN: feature-guiding generative adversarial networks for text generation, *IEEE Access* 8 (2020) 105217–105225.
- [59] Q. Wu, L. Li, Z. Yu, TextGAIL: generative adversarial imitation learning for text generation, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, AAAI Press, 2021.
- [60] B.D. Ziebart, A.L. Maas, J.A. Bagnell, A.K. Dey, Maximum entropy inverse reinforcement learning, in: 23rd National Conference on Artificial Intelligence, vol. 3, 2008, pp. 1433–1438.
- [61] O. Nachum, M. Norouzi, K. Xu, D. Schuurmans, Bridging the gap between value and policy based reinforcement learning, in: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, 2017, pp. 2772–2782.
- [62] I. Sutskever, O. Vinyals, Q.V. Le, Sequence to sequence learning with neural networks, in: *Advances in Neural Information Processing Systems*, 2014, pp. 3104–3112.
- [63] Y. Zhang, Z. Gan, L. Carin, Generating text via adversarial training, *NIPS Workshop on Adversarial Training*, vol. 21, 2016.
- [64] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, in: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, Minneapolis, Minnesota, 2019, pp. 4171–4186.
- [65] K. Clark, M.-T. Luong, Q.V. Le, C.D. Manning, Electra: pre-training text encoders as discriminators rather than generators, in: 8th International Conference on Learning Representations, 2020.
- [66] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, Language models are unsupervised multitask learners, *OpenAI Blog* 1 (8) (2019) 1–9.

Gustavo Henrique de Rosa Bachelor in Computer Science at São Paulo State University (2016), focusing on pattern recognition, machine learning algorithms, and meta-heuristic optimization. Master of Science in Computer Science at São Paulo State University (2018), focusing on deep learning and meta-heuristic optimization. Currently, Ph.D. student in Computer Science at São Paulo State University.

João Paulo Papa B.Sc. in Information Systems from São Paulo State University. In 2005, received his M.Sc. in Computer Science from Federal University of São Carlos. In 2008, received his Ph.D. in Computer Science from University of Campinas. He has been a Professor at Computer Science Department, São Paulo State University since 2009.