



A novel individual-relational consistency for bad semi-supervised generative adversarial networks (IRC-BSGAN) in image classification and synthesis

Mohammad Saber Iraj¹ · Jafar Tanha¹ · Mohammad-Ali Balafar¹ · Mohammad-Reza Feizi-Derakhshi¹

Accepted: 14 July 2024 / Published online: 6 August 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Semi-supervised learning leverages both labeled and unlabeled images for model training, addressing the scarcity of labeled data. However, challenges persist, including the determination of appropriate thresholds for pseudo-labeling, the effective utilization of uncertain unlabeled images, the absence of consistency regularization, and the oversight of inter-image relationships among images in low-density areas. This study introduces a novel approach named the Individual-Relational Consistency for Bad Semi-supervised Generative Adversarial Networks (IRC-BSGAN) to tackle these issues. IRC-BSGAN integrates bad adversarial training, consistency regularization, and pseudo-labeling to reduce error rates and enhance classifier performance. It includes various components, such as a bad generator network, a discriminator network, a classifier, and consistency regularization modules. IRC-BSGAN introduces new individual and relational consistency regularization losses on bad fake images in low-density areas, thereby generating informative images that precisely estimate the classifier's decision boundary. The proposed method ensures diversity and consistent labeling of bad fake images by integrating consistency mechanisms. It particularly focuses on low-density areas and extracts extra semantic details from these images by promoting local consistency and coherence among them. The effectiveness of IRC-BSGAN is realized by improving the pseudo-labeling of unlabeled images, especially for low-confidence unlabeled images. For the SVHN dataset with 1000 labeled training images and the CIFAR-10 dataset with 4000 labeled training images, the error rate reduced from 3.89 to 3.67 and from 7.29 to 6.17, respectively. Similarly, on the CINIC-10 dataset with 1000 labeled training images per class, IRC-BSGAN achieved a reduction in error rate from 19.38 to 15.45. On the COVID-19 dataset with 30 labeled training images, the error rate decreased from 7.41 to 5.55.

Keywords Semi-supervised classification · Informative fake images · Low-confidence images · Individual consistency regularization · Relational consistency regularization

1 Introduction

Supervised learning, heavily reliant on labeled images, has historically dominated machine learning [1]. However, the manual labeling process is expensive, time-consuming, and impractical for large datasets [2]. The scarcity of labeled images in real-world scenarios presents significant challenges to supervised learning's effectiveness. Semi-supervised learning has emerged as a promising alternative, combining limited labeled data with extensive unlabeled data across various domains [3]. Techniques such as pseudo-labeling [4] and consistency regularization [5, 6] have shown promise in enhancing classification model performance by uncovering additional patterns from unlabeled data. One potential line

✉ Jafar Tanha
tanha@tabrizu.ac.ir; jafar.tanha.pnu@gmail.com

Mohammad Saber Iraj
iraji.ms@gmail.com; iraji.ms@pnu.ac.ir

Mohammad-Ali Balafar
Balafarila@tabrizu.ac.ir

Mohammad-Reza Feizi-Derakhshi
mfeizi@tabrizu.ac.ir

¹ Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

of questioning is how can these semi-supervised learning techniques effectively harness unlabeled data to address the limitations of supervised learning in contexts where labeled data is scarce or unavailable.

Pseudo-labeling in semi-supervised classification involves generating pseudo-labels for unlabeled images using limited labeled images [7]. The role of unlabeled images and pseudo-labels is crucial in semi-supervised learning. Algorithms assign pseudo-labels based on high predicted probabilities or threshold-exceeding probabilities [8, 9]. Nonetheless, using fixed thresholds limits the effective use of unlabeled images, making it challenging to determine suitable thresholds for improving classifier performance [10]. Thus, effectively utilizing low-confidence unlabeled images below the threshold remains difficult [11]. Additionally, addressing incorrect pseudo-labels and their impact on consistency regularization techniques remains an ongoing challenge.

Consistency regularization proves valuable in assigning pseudo-labels to unlabeled images below a specified threshold, particularly when they are proximate to the decision boundary [12]. This method ensures the model consistently assigns class labels even amid various alterations to inputs [13]. Local regularization seeks to harmonize labels between unlabeled and labeled images at a local level, often through augmentations and perturbations to alleviate data sparsity [14]. These augmentations might involve minor adjustments such as rotation and translation for images or the introduction of adversarial noise via techniques like virtual adversarial training [15, 16]. Nevertheless, applying local alterations to images close to the decision boundary may push them beyond the correct class label boundary, diminishing the efficacy of local consistency regularization and impeding learning efficiency.

Label consistency regularization combines weak and strong augmentation techniques for an accessible image, utilizing augmentation anchoring and pseudo-labeling with a probability surpassing a predefined threshold (e.g., the Fix-Match method [17]). However, challenges arise, including determining prediction probability using a fixed threshold and generating robust image augmentation [18]. Dynamic thresholding can address these issues by selecting additional pseudo-labels for classes, resulting in partial learning improvements [19, 20]. Nevertheless, caution is warranted when employing methods like FlexMatch [4] and FullMatch [10] with extremely low probability thresholds, as they may hinder efficient learning due to pseudo-label uncertainty.

One limitation of current consistency-based methods is their neglect of the connections among samples in low-confidence unlabeled images [21]. The inherent relationships among images, such as labels, have been shown to significantly enhance their informativeness [22, 23]. However, existing methods focus narrowly on perturbing individual data points and overlook the connections and relationships

among different images. Consequently, they fail to fully utilize the valuable information naturally present in the structural relationships among low-confidence, unlabeled images [24]. This limitation can lead to a situation where the function displays smoothness near each unlabeled image in low-density areas but lacks smoothness among them.

Recent advances in generative adversarial networks (GANs) [25] have demonstrated their effectiveness in improving generalization for semi-supervised learning by moving the decision boundary towards low-density areas [26, 27]. One approach, known as bad generative adversarial networks, seeks to improve a poorly performing generator that produces fake images near the true classifier boundary [28]. These generated images, similar to decision support vectors, help the classifier extract discriminative information between classes and enhance generalization performance [29]. Alternatively, marginal generative adversarial networks propose a three-player game involving a bad generator instead of the traditional two-player game [30]. Nevertheless, bad generative adversarial networks face challenges due to the lack of effective label consistency regularization, which results in inconsistent labels for bad fake samples. As far as we know, no specific proposal for individual-relational consistency regularization has been made to address the issues associated with bad GANs.

To tackle these challenges, we investigate individual and relational consistency among images in low-density areas. Our goal is to promote local smoothing of each bad fake image and smoothness among similar neighboring points in low-density regions. This approach aims to extract more semantic information from low-confidence unlabeled images, leading to improved performance. We introduce a novel method called Individual-Relational Consistency for Bad Semi-supervised Generative Adversarial Networks (IRC-BSGAN). This method aims to enhance semi-supervised classification in low-density areas by leveraging labeled and unlabeled images below a specific probability threshold. The framework includes a bad generator, a discriminator, and a classifier. The bad generator generates images near the correct decision boundary of the classifier. Simultaneously, individual and relational consistency regularization improves the performance of the bad generator and discriminator. Relational consistency provides more structural and semantic information for the bad fake images compared to individual consistency by assessing the relationship between one bad image and the others. Experimental results on the SVHN, CINIC-10, COVID-19, and CIFAR-10 datasets show that our proposed semi-supervised model outperforms existing approaches in the literature. The key contributions of our research are outlined below:

- 1- We present a novel semi-supervised model featuring a bad generator, integrating individual and relational con-

sistency regularization between latent data and bad fake images in low-density regions.

- 2- We introduce novel individual consistency regularization losses on bad fake images to improve their generation and accurately predict pseudo-labels for low-confidence unlabeled images.
- 3- We propose novel (inversed) relational consistency regularization losses operating on the latent vectors of bad fake images in low-density areas. This regularization technique aims to improve coherence and consistency within the latent space and the bad generated images.

This study is structured as follows: Section 1 provides an introduction. Section 2 reviews relevant literature. Section 3 outlines the proposed model. The experimental results are presented in Section 4. Section 5 comprises the discussion. Lastly, Section 6 concludes the study and discusses future works. For interested parties, the Python code related to our research can be accessed at <https://github.com/ms-iraji/IRCRBSGAN>.

2 Related works

In this section, we examine the existing literature and research pertaining to consistency regularization and semi-supervised classification, which are relevant to our study. Self-learning, as an initial approach in semi-supervised learning, entails predicting pseudo-labels for unlabeled images based on the knowledge gleaned from labeled images [31]. In this method, unlabeled images within the training dataset serve as training data and provide additional information to the model when their pseudo-labels exceed a threshold of 0.95 during each epoch of model training. However, the self-learning method discards information from unlabeled images below the threshold [32].

The teacher-student method is a widely used framework for consistency regularization [33]. In this approach, the model serves as a teacher, generating labels, while also functioning as a student by learning from its own generated labels. Label consistency regularization is applied to the labels produced by both the teacher and student networks using weight perturbations. However, the efficiency of this method is notably diminished because of the possibility of the teacher generating incorrect pseudo-labels, which the student then learns.

Li et al. introduced the SRC-MT architecture for semi-supervised learning with unlabeled images [22]. It comprises a teacher model and a student model, both utilizing a deep neural network (DNN). These models enforce consistency regularization by aligning pseudo-labels under weight perturbations and input noise. Relational consistency

regularization enhances learning by encouraging similar activations for related unlabeled images. However, a weakness arises in local label smoothing near the decision boundary without adequate augmentation anchoring, which may lead to incorrect pseudo-labels.

The Π model conducted two forward propagations on unlabeled images in each training cycle [34]. During these propagations, random data perturbations and dropouts were applied to the images [35]. The model aimed for the forward propagation to produce identical predictions for two labels. Initially, the Π model required sending images twice during each training iteration. Thus, the temporal ensemble model (TE) was introduced to alleviate this overhead. The TE model utilizes an exponential moving average technique to aggregate class-label predictions.

The TE model [36] incorporates random augmentations, such as the dropout layer, for unlabeled images in a neural network. In this method, consistency regularization is implemented by comparing predictions for one augmentation of an image with the exponential moving average of predictions for another augmentation of the same image [35]. Nevertheless, a drawback of this approach is the time-consuming adjustment of predictions after training on a large dataset.

The mean teacher approach (MT) demonstrates the importance of averaging model weights rather than solely relying on label predictions [37]. It establishes a teacher model by averaging weights from consecutive student models and integrates exponential moving average weights from the student model [37]. This amalgamation of information enhances target labels and intermediate representations, thereby improving learning and performance. The MT method offers advantages such as a quicker feedback loop between student and teacher models and scalability for large datasets [38].

DSSLDDR [39] is a semi-supervised learning model designed to address the scarcity of labeled images by combining dictionary representation with deep learning. It leverages class-specific dictionaries to reconstruct images and extract discriminative features using a deep neural network. Additionally, an entropy regularization term is utilized to handle unlabeled images. To improve the class estimation accuracy, DSSLDDR + incorporates consistency/contrastive learning. However, a limitation of this approach is its constrained integration of dictionary learning, which restricts the potential advantages across all layers of the model.

The authors [40] introduced a novel graph-based semi-supervised learning method named dynamic anchor graph embedding (DAGE), with the goal of embedding graphs and classifying sample nodes simultaneously. DAGE utilizes a dynamic anchor graph constructed within the

latent space of a neural network to improve graph quality and simplify the embedding learning problem. It employs a two-branch network architecture, comprising a dynamic graph embedding branch and a single-sample consistency branch, to integrate local and global information, resulting in more reliable classification outcomes. DAGE distinguishes itself by continuously optimizing the graph structure and integrating graph structural and model discriminant information derived from both labeled and unlabeled data. However, like other semi-supervised approaches, DAGE's effectiveness depends on the assumption that data exhibits cluster structures or resides on smooth manifolds, which may pose limitations in real-world applications.

In the study [8], researchers introduced DNLL, a novel semi-supervised classification framework called dual pseudo-negative label learning. This framework employs two sub-models to generate pseudo-negative labels and includes a selection mechanism based on estimating uncertainty. By enhancing the utilization of unlabeled images and reducing model parameter coupling, the framework achieves enhanced performance and generalization. However, limitations of the approach include the quality of the pseudo-negative labels and the potential reliance on specific selection criteria.

In the paper [41], ReliaMatch is introduced as a semi-supervised classification approach addressing the challenges associated with utilizing unlabeled data and handling error information. ReliaMatch effectively integrates various techniques, including confidence thresholding, curriculum learning, pseudo-label filtering, and feature filtering, to notably enhance the accuracy and reliability of classification tasks. By employing the confidence threshold, the algorithm filters out unreliable information, eliminates ambiguous semantic features, and discards unreliable pseudo-labels. Nonetheless, it is crucial to note that the algorithm's performance hinges on the accurate calibration and control of the confidence threshold. Additionally, the choice of feature extraction methods and model architectures can significantly impact its overall performance.

The authors introduced collaborative learning with unreliability adaptation (CoUA), an approach for semi-supervised image classification that emphasizes cooperation among multiple networks [42]. CoUA enables the networks to adjust their predictions and establish customized training objectives for unlabeled data. By integrating an adaptation module, the networks exchange training experiences and learn transition probabilities between their predictions. This collaborative framework effectively addresses the challenge of unreliable predictions by fostering collaboration while minimizing negative effects. The approach promotes

consistent predictions and resilience against adversarial perturbations, thereby enhancing the collaborative learning process. However, the research emphasises the importance of exploring training experience exchange among all networks and integrating predictions with uncertainty clues as areas for further enhancement.

In [43], a novel semi-supervised GAN architecture named Triple-BigGAN is introduced, which integrates generative modeling and classification tasks. Triple-BigGAN extends the BigGAN network to learn a discriminative classifier while generating high-quality synthesized images using partially labeled data. The framework comprises a discriminator, classifier, and generator. The classifier is trained on both real labeled data and generated samples. The discriminator distinguishes real image-label pairs from the labeled dataset and pairs obtained from the classifier for the unlabeled dataset. Unlike previous frameworks such as Triple-GAN and EC-GAN, Triple-BigGAN prioritizes the end-to-end training of a robust classifier and discriminator. The primary goal of Triple-BigGAN is to achieve semi-supervised joint distribution matching, enabling the utilization of labeled and unlabeled data for image classification and synthesis within a good generator.

Virtual adversarial training (VAT) [44] was developed to regularize the distribution of conditional labels surrounding a particular input, safeguarding against local disruptions. The objective is to preserve the original image's label even in the presence of localized perturbations around each image [45]. However, during local augmentation of the image near the decision boundary, it inadvertently shifts the image to the opposite side of the class boundary [46]. Consequently, this approach may not be as effective in points located near the correct class boundary, potentially impacting its efficiency.

The FixMatch method [47] integrates supervised and unsupervised learning with augmentation anchoring. Weakly augmented labeled images are utilized to facilitate supervised learning. Semi-supervised learning involves the conversion of predicted labels exceeding a specific threshold of weak augmentations to strong augmentations of the same image. Weak augmentation involves basic transformations, while strong augmentation encompasses multiple transformations. The method excludes images with low label prediction probabilities, aiding in establishing the decision boundary.

Interpolation consistency training (ICT) is a semi-supervised learning method that employs a mixup technique to shift the decision boundary away from class boundaries [5]. Mixup entails interpolating pairs of images, promoting straightforward linear behavior and improving generalization [48]. It diminishes label

memorization and bolsters resilience against adversarial images. However, when applied to low-confidence images, ICT faces limitations.

The authors suggested employing a generator to produce bad fake images closely resembling real ones [26]. These artificially generated samples positioned the discriminator boundary between distinct image classes, thereby reducing its generalization error. The margin generative adversarial network (margin GAN) [30] was devised as a three-player system to generate substandard images in low-density regions. The discriminator distinguished real images from generator-generated ones, while the classifier aimed to maximize the margin of real images and decrease the margin of fake ones. Nonetheless, consistency regularization can still enhance semi-supervised learning across bad GAN models in low-density areas.

3 Proposed method

In this section, the IRC-BSGAN approach is introduced as a three-player adversarial game comprising a discriminator, a classifier, and a bad generator. The framework integrates latent individual-relational consistency regularization to address issues with bad fake images. Furthermore, we propose latent inversed individual-relational consistency regularization for the bad semi-supervised generative adversarial model. Similar to bad GAN, the IRC-BSGAN utilizes adversarial training and incorporates redesigned elements, including loss functions and relational consistency, in order to improve its performance. The model architecture is depicted in Fig. 1, with the proposed workflow and algorithm outlined in Fig. 2 and Algorithm 1, respectively.

3.1 Bad generator

The limited availability of data presents a significant challenge in semi-supervised generative learning. A good generator assumes a pivotal role in generating high-quality images and mitigating the constraints imposed by limited data resources [43]. In another approach, the generator conditions class information to derive the label sample distribution [49]. Recent studies [26, 29, 30] have increasingly emphasized the integration of bad generator modules into semi-supervised generative learning alongside conventional good generators. These bad generator modules are purposefully designed to produce examples aiding the classifier or discriminator network in delineating decision boundaries. Although two and three-player game approaches based on the bad generator have been

proposed, they encounter challenges such as mode collapse, absence of label smoothing, and inconsistency. We formulate the details of the basic bad generator as follows.

The generator is tasked with learning the distribution of real images and generating fake images that closely resemble realistic ones from the perspective of the discriminator. It is implemented as a deep neural network (DNN) composed of transposed convolutional layers [50]. Utilizing a latent vector $z \sim p_z$, which is a stochastic noise vector sampled from a distribution (typically either normal $\mathcal{N}(0,1)$ or Gaussian), the generator produces fake images [49, 51, 52]. These bad images are then presented to the discriminator as real (labeled 1) during the adversarial game, as shown in Eq. 1.

$$\text{Loss}_{\text{generator-adversarial-base}} = - \sum_{z \sim p_z} \log D(G(z)) \quad (1)$$

In a separate adversarial game involving the classifier, the bad generator generates images where the classifier has a high margin [53]. Subsequently, the classifier's prediction loss for the images generated by the bad generator, specifically with the class with the highest probability as the target label, is calculated [54]. The generator's parameters are then updated using the cross-entropy function (Eq. 2).

$$\text{Loss}_{\text{generator-adversarial-2}} = - \sum_{i=1}^{nc} (\arg \max(C(G(z))_i) \log(C(G(z))_i)) \quad (2)$$

nc = number of class

3.2 Discriminator

The discriminator network serves a vital function in generative adversarial networks. Recently, a technique leveraging clustering, an augmented feature matching strategy, and multiple sub-discriminators has been integrated into the two-player bad GAN configuration [26]. We aim to augment classification accuracy and mitigate mode collapse. The discriminator helps train the generator by providing feedback on the quality of the generated samples. Nevertheless, in this study, we implement individual-relational regularization to further enhance the performance of the discriminator network. The specifics of a fundamental discriminator are elucidated in the subsequent subsection.

In the adversarial game involving the generator G , the discriminator D is a DNN responsible for discerning between real and fake generated images. For discriminator training, the real labeled images $x^l \sim p_{x^l}^{\text{data}}$ and unlabeled images $x^u \sim p_{x^u}^{\text{data}}$ are assigned the label "1", while the fake-generated images $x^g = G(z) \sim p_{x^g}^{\text{fake}}$ are labeled as "0". This distinction

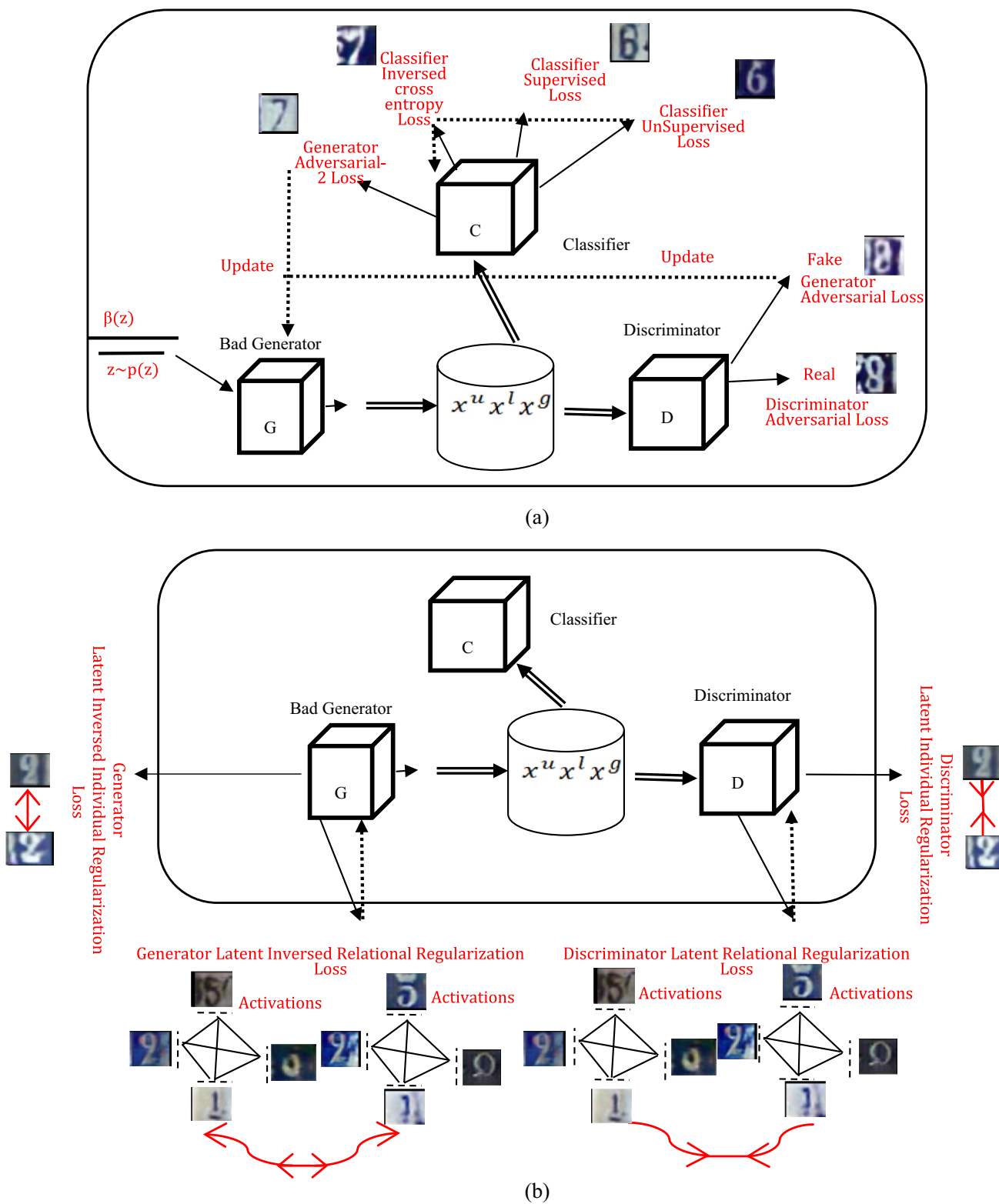


Fig. 1 Architectural overview of the IRC-BSGAN; (a) 3-player bad GAN, (b) 3-player bad GAN with individual-relational consistency

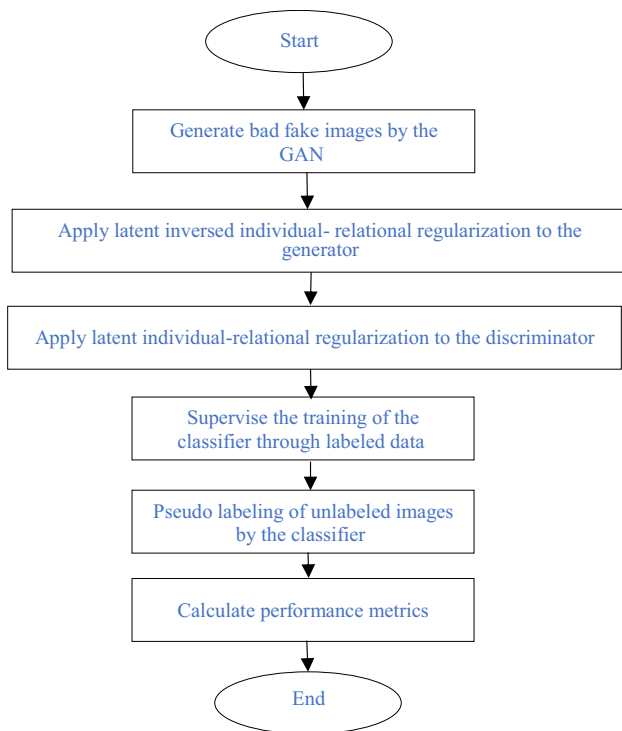


Fig. 2 The overall workflow of individual-relational consistency for the bad semi-supervised generative adversarial networks (IRC-BSGAN) method

is achieved using the base adversarial GAN function [55], which is applied to the discriminator through Eq. 3.

$$\text{Loss}_{\text{discriminator-adversarial}} = - \sum_{x^l, x^u \sim p^{\text{data}}} [\log D(x)] - \sum_{x \sim p^{\text{fake}}} [\log(1 - D(x))] \quad (3)$$

3.3 Classifier

The classifier plays a pivotal role in semi-supervised learning by assigning class labels to labeled, unlabeled, and fake images. However, incorrect pseudo-labeling of unlabeled images can significantly impair the model's efficacy and result in inaccurate decision boundaries. Recent approaches [43, 56] have introduced the classifier as a separate element alongside the traditional discriminator, which was tasked with assigning class labels to images. In subsequent subsections, novel techniques are applied to bad GANs to strengthen the classifier and improve class estimation accuracy. This subsection outlines the formulation of the classifier within a bad GAN framework.

A DNN-based multi-class classifier, denoted as C , is utilized for data label prediction. In supervised learning, where

the real labeled image $(x^l, y) \sim p_{(x^l, y)}^{\text{data}}$ is available, the classifier aims to minimize the error between the predicted label $C(x^l)$ and the true label y using a cross-entropy function (Eq. 4) [57, 58].

$$\text{Loss}_{\text{classifier-real-supervised}} = - \sum_{i=1}^{nc} (y_i \log(C(x^l)_i)) \quad (4)$$

In the adversarial game between the generator and the classifier, the classifier assigns low-margin labels to the bad-generated images utilizing an inverted cross-entropy loss function (Eq. 5). This adversarial game between the classifier and generator results in the production of bad fake images that exhibit similarities to the support vectors containing decision boundary information of the classifier.

$$\text{Loss}_{\text{classifier-inversed-cross-entropy}} = - \sum_{i=1}^{nc} (\arg \max(C(G(z))_i) \log(1 - C(G(z))_i)) \quad (5)$$

Given the scarcity of labeled images, the classifier leverages a large number of unlabeled images $x^u \sim p_{x^u}^{\text{data}}$, utilizing the pseudo-labeling loss function. The predicted label $C(x^u)$ tends to approximate the class with the maximum probability [59], represented as $\arg \max(C(x^u))$, serving as the target label for unlabeled images (Eq. 6).

$$\text{Loss}_{\text{classifier-real-unsupervised}} = - \sum_{i=1}^{nc} (\arg \max(C(x^u)_i) \log(C(x^u)_i)) \quad (6)$$

$nc = \text{number of class}$

3.4 Consistency regularization modules

Many methods, such as [5, 22, 23, 35, 37, 41, 43, 60], implement consistency regularization techniques on unlabeled images, facing challenges, particularly with low-confidence images near decision boundaries. In our proposed approach, we introduce innovative consistency regularization terms on bad fake images utilizing 3-player bad GANs to tackle these challenges directly. These terms are engineered to bolster the treatment of low-confidence images proximate to decision boundaries.

The proposed consistency regularization modules are integral to the recommended semi-supervised GANs, aiming to elevate the performance and robustness of the model components. These modules consist of two discrete submodules: latent individual-relational consistency regularization and latent inverse individual-relational consistency regularization. Through combining these two consistency regularization submodules, both the discriminator and the bad generator's performance is bolstered, thereby augmenting the generation of bad fake images as support vectors. This improvement leads to enhanced

generalization and performance of the suggested semi-supervised model.

3.4.1 Latent individual-relational regularization

To boost the discriminator's performance and counter the effects of limited image availability, we employ augmentation and consistency regularization techniques. In individual

consistency regularization, we label bad-generated images consistently as fake, whether they result from latent space modifications or local alterations within the same latent space (Eq. 7). To introduce local variations to each image's latent vectors, we incorporate a vector of random numbers using a β function with mean μ and variance σ . This approach significantly reduces the label gap between bad images generated from the latent space and those resulting from local changes [61].

$$\text{Loss}_{\text{discriminator-latent-individual-regularization}} = \left\| \sum_{z \sim P_z} D(G(z)) - \sum_{z' \sim P_{\beta(z)}} D(G(z')) \right\|_2^2 \tag{7}$$

Relational consistency regularization considers the inherent relationship between a bad fake image and other instances [22], capturing additional semantic information. We extract image activations (IA) from the fully connected layer of the discriminator D_f to serve as image features. Subsequently, we compute the activations of a batch of bad generated images originating from the latent space and its locally modified versions. Then, the image relationship matrix RL is computed by internally multiplying the activations of different images within the mini-batch using Eqs. 8 and 9.

$$\text{RL}_{i,j} = \left[\frac{\text{IA}_i \cdot (\text{IA}_j)^T}{\|\text{IA}_i \cdot (\text{IA}_j)^T\|_2} \right], \forall i, j = 1 \dots \text{bz}, \text{bz} = \text{batch size}, \text{T} = \text{transpose operation} \tag{8}$$

$$D_f(x_i) = \text{IA}_i, \forall i = 1 \dots \text{bz} \tag{9}$$

We utilize latent relational consistency regularization to preserve the structural stability of the relationship between bad images generated from the latent space when subjected to local perturbations of the same latent vector. This regularization is integrated into the discriminator and is represented by Eq. 10.

$$\text{Loss}_{\text{discriminator-latent-individual-regularization}} = \left\| \sum_{z \sim P_z} \text{RL}(D_f(G(z))) - \sum_{z' \sim P_{\beta(z)}} \text{RL}(D_f(G(z'))) \right\|_2^2 \tag{10}$$

3.4.2 Latent inverse individual-relational regularization

In the individual consistency regularization loss function (Eq. 11), the bad generator, in contrast to the discriminator, generates diverse bad images by inversely ensuring consistency between the latent space vector and its local perturbations. Furthermore, an inverse relational consistency regularization loss function is introduced to disrupt the preservation of structural relationships between the activations of bad images from the latent space $G(z)$ under

local perturbations of the same latent vector within a mini-batch (Eq. 12). This methodology enhances the generator's ability to glean discriminative information from the latent space and translate it into a broader spectrum of bad fake images, thereby augmenting both variety and sensitivity within the latent space. Moreover, it prevents the generator from producing similar structural patterns for similar latent vectors, fostering exploration of the latent space. Figure 2 illustrates the overall flow, while Algorithm 1 provides the pseudo-code implementation of the proposed method.

Algorithm 1 Individual-relational consistency for bad semi-supervised generative adversarial networks (IRC-BSGAN).

Inputs: $(x^l, y) \sim p_{(x^l, y)}^{\text{data}}$: input-label pair, $x^l \sim p_{x^l}^{\text{data}}$: real labeled images, $x^u \sim p_{x^u}^{\text{data}}$: real unlabeled images, $z \sim p_z$: latent vector, β : noise function, θ_G : generator parameters, θ_D : discriminator parameters, θ_C : classifier parameters

For $i = 1$ of the number of training iterations, do the following:

Sample batch $z \sim p_z, x^l \sim p_{x^l}^{\text{data}}, x^u \sim p_{x^u}^{\text{data}}$

➤ Create a weakly augmented latent vector

Sample batch $z' \sim p_{\beta(z)}$

➤ Generate fake images $x^g \sim p_{x^g}^{\text{fake}}$

$\text{Loss}_{\text{generator-adversarial-base}} \leftarrow \sum_{z \sim p_z} \log D(G(z)), x^g \leftarrow G(z)$

➤ Increase the margin on the generated fake images

$\text{Loss}_{\text{generator-adversarial-2}} \leftarrow -\sum_{i=1}^{nc} (\arg \max(C(G(z)))_i) \log(C(G(z)))_i$

➤ Apply latent inversed individual consistency regularization to G

$\text{Loss}_{\text{generator-latent-inversed-individual-regularization}} \leftarrow -\left\| \sum_{z \sim p_z} D(G(z)) - \sum_{z' \sim p_{\beta(z)}} D(G(z')) \right\|_2^2$

➤ Apply latent inversed relational consistency regularization to G

$\text{Loss}_{\text{generator-latent-inversed-relational-regularization}} \leftarrow -\left\| \sum_{z \sim p_z} RL(D_f(G(z))) - \sum_{z' \sim p_{\beta(z)}} RL(D_f(G(z'))) \right\|_2^2$

Compute generator losses and update its parameters θ_G via the Adam optimizer.

➤ Distinguish between real and fake images

$\text{Loss}_{\text{discriminator-adversarial}} \leftarrow -\sum_{x^l, x^u \sim p^{\text{data}}} [\log D(x)] - \sum_{x \sim p^{\text{fake}}} [\log(1 - D(x))]$

➤ Apply latent individual consistency regularization to D

$\text{Loss}_{\text{discriminator-latent-individual-regularization}} \leftarrow \left\| \sum_{z \sim p_z} D(G(z)) - \sum_{z' \sim p_{\beta(z)}} D(G(z')) \right\|_2^2$

➤ Apply latent relational consistency regularization to D

$\text{Loss}_{\text{discriminator-latent-relational-regularization}} \leftarrow \left\| \sum_{z \sim p_z} RL(D_f(G(z))) - \sum_{z' \sim p_{\beta(z)}} RL(D_f(G(z'))) \right\|_2^2$

Compute discriminator losses and update the parameters θ_D via the Adam optimizer.

➤ Supervise the training of the classifier using cross-entropy

$\text{Loss}_{\text{classifier-real-supervised}} \leftarrow -\sum_{i=1}^{nc} (y_i \log(C(x^l)_i))$

➤ Reduce the margin through inversed cross-entropy

$\text{Loss}_{\text{classifier-inversed-cross-entropy}} \leftarrow -\sum_{i=1}^{nc} (\arg \max(C(G(z)))_i) \log(1 - C(G(z)))_i$

➤ Increase the margin on the unlabeled images

$\text{Loss}_{\text{classifier-real-unsupervised}} \leftarrow -\sum_{i=1}^{nc} (\arg \max(C(x^u)_i) \log(C(x^u)_i))$

Compute classifier losses and update the parameters θ_C via the Adam optimizer.

$$\text{Loss}_{\text{generator-latent-inversed-individual-regularization}} = -\left\| \sum_{z \sim p_z} D(G(z)) - \sum_{z' \sim p_{\beta(z)}} D(G(z')) \right\|_2^2 \quad (11)$$

$$\text{Loss}_{\text{generator-latent-inversed-individual-regularization}} = -\left\| \sum_{z \sim p_z} RL(D_f(G(z))) - \sum_{z' \sim p_{\beta(z)}} RL(D_f(G(z'))) \right\|_2^2 \quad (12)$$

Table 1 Confusion matrix

		Actual	
		Negative	Positive
Predicted	Negative	True Negatives	False Negatives
	Positive	False Positives	True Positives

Table 2 The proposed model parameters

Parameters	SVHN	CIFAR-10	CINIC-10/ COVID-19
lrD	0.0002	0.0002	0.0002
lrG	0.0002	0.0002	0.0002
lrC	0.1	0.1	0.1
Momentum	0.5	0.5	0.9
Beta1	0.5	0.5	0.5
Beta2	0.999	0.999	0.999
batch_size	100	100	128
β function:(μ, σ)	(0,0.03)	(0,0.07)	(0,1)
z_dim	100	100	100
epochs	400	400	400

4 Experiments and results

In this section, we conduct experiments to assess our proposed semi-supervised model's effectiveness. We train our model using a limited amount of labeled data and a large pool of unlabeled images to build a classifier across four datasets. Subsequently, we evaluate the classifier's performance on the testing set. Details about the datasets are provided in Section 4.1, while Section 4.2 defines performance metrics. We present the network architecture and parameter setup in Section 4.3. Finally, Section 4.4 documents the outcomes of implementing our proposed method.

4.1 Datasets

The efficiency of the proposed semi-supervised model was evaluated on images from SVHN [62], CINIC-10 [39], COVID-19 [63], and CIFAR-10 [64] data sets.

- The SVHN (Street View House Numbers) dataset comprises 73,257 training images and 26,032 test images. Each sample is a 32×32 color image displaying cropped house numbers ranging from 0 to 9 against various backgrounds.
- The CIFAR-10 (Canadian Institute for Advanced Research, 10 classes) dataset contains 50,000 training and 10,000 test images. Each image is a 32×32 color image representing one of 10 classes of natural objects: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck.
- CINIC-10 serves as a bridge between CIFAR-10 and ImageNet, comprising 270,000 RGB images (32×32

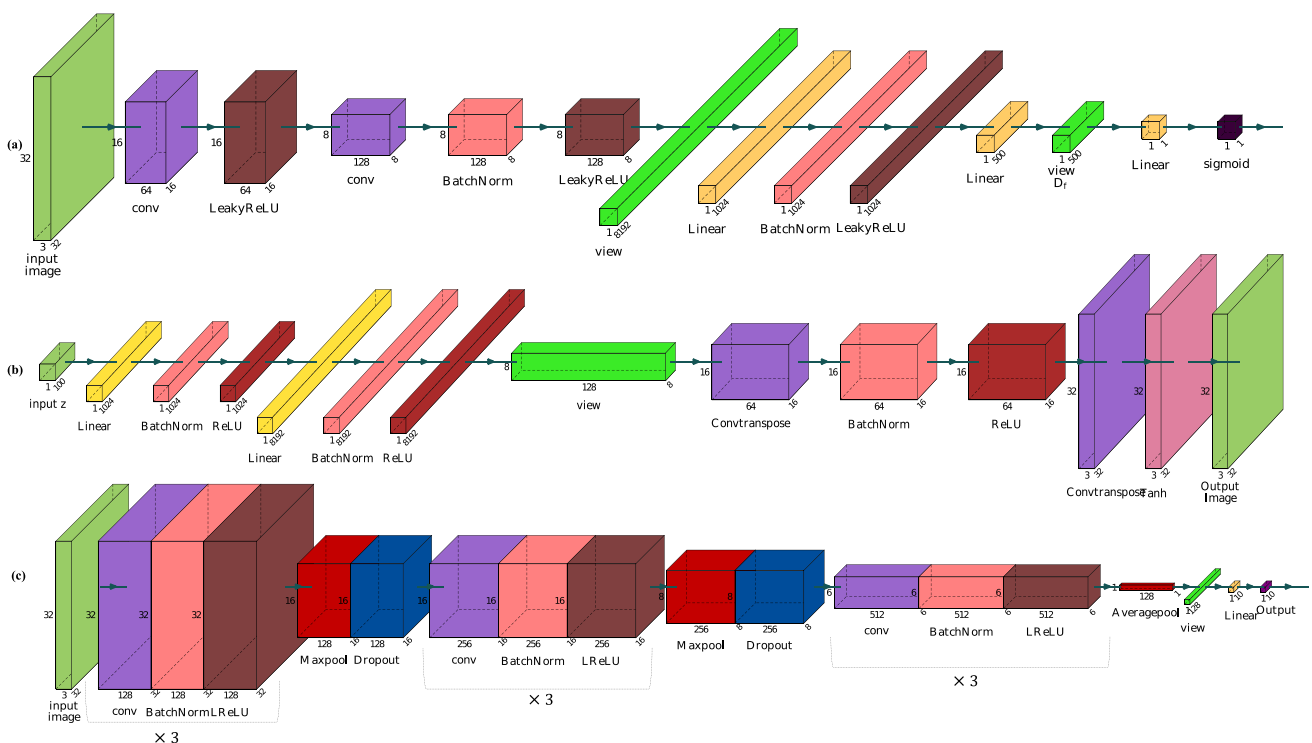


Fig. 3 The structure of the a) discriminator, b) generator, and c) classifier for SVHN, CINIC-10, COVID-19, and CIFAR-10

Fig. 4 Confusion matrices for classifiers trained with (a) 500 and (b) 1000 labeled training images on SVHN test data

Actual Label	0	1	2	3	4	5	6	7	8	9
0	1689	8	5	5	0	4	10	7	2	5
1	20	4786	16	20	28	6	7	207	6	3
2	6	22	4039	29	9	4	5	20	5	10
3	4	21	11	2772	2	19	3	4	18	28
4	3	17	6	7	2467	0	1	4	1	17
5	6	9	6	37	4	2280	27	3	7	5
6	10	8	1	3	7	10	1931	1	4	2
7	1	71	12	4	6	2	0	1918	2	3
8	8	5	2	6	6	2	38	0	1578	15
9	20	7	24	5	3	3	2	0	3	1528
	0	1	2	3	4	5	6	7	8	9

(a)

Actual Label	0	1	2	3	4	5	6	7	8	9
0	1693	9	6	6	0	7	8	3	1	11
1	20	4859	23	17	42	6	6	114	9	3
2	6	24	4060	12	14	4	4	15	4	6
3	6	19	11	2750	7	14	7	4	18	46
4	3	19	4	6	2464	3	3	4	1	16
5	8	7	5	29	6	2289	25	3	5	7
6	11	10	2	4	4	7	1930	1	5	3
7	4	65	12	7	6	2	0	1920	1	2
8	14	5	1	11	7	2	27	1	1576	16
9	21	6	23	4	2	3	2	0	3	1531
	0	1	2	3	4	5	6	7	8	9

(b)

pixels). It facilitates the evaluation of model performance on ImageNet-like images corresponding to the classes in CIFAR-10. The dataset shares class labels with CIFAR-10 and is segmented into training, validation, and test subsets, each containing 90,000 images.

- Canayaz developed a comprehensive COVID-19 X-ray dataset featuring three distinct patient subgroups: COVID-19, pneumonia, and healthy individuals. The dataset comprises a total of 1092 images, with each class being balanced. All grayscale images in the experiment are 32×32 pixels. For our study, we reserved a subset of 108 images for testing, with each class containing 36 images.

The training datasets comprise real images, whereas the images produced by the generator are deemed fake. In semi-supervised learning, a subset of the training images, accompanied by their respective labels, are identified as labeled real images. Conversely, the remaining training images lacking labels are denoted as unlabeled real images.

4.2 Evaluation metric

Table 1 presents the confusion matrix, consisting of the components TP (true positives), TN (true negatives), FP (false positives), and FN (false negatives) [65]. In multi-class classification scenarios, the positive class is

identified as the target class, while the remaining classes are treated as negative. TP signifies the number of actual positive class samples correctly classified as positive, while TN denotes the number of actual negative class samples correctly classified as negative. FP represents the count of actual negative samples incorrectly classified as positive, whereas FN indicates the number of actual positive samples incorrectly classified as negative. The accuracy index is computed using Eq. 13, where a higher value signifies superior model performance [66]. Additionally, the model's error rate, F1 score, sensitivity, specificity, and precision metrics are defined by Eqs. 14–18.

$$\text{Accuracy} = ((\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{FN} + \text{TN})) \times 100 \quad (13)$$

$$\text{Error rate} = 1 - \text{Accuracy} \quad (14)$$

$$\text{F1} = ((2 \times \text{TP}) / ((2 \times \text{TP}) + \text{FP} + \text{FN})) \quad (15)$$

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN}) \quad (16)$$

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP}) \quad (17)$$

$$\text{Precision} = \text{Tp} / (\text{Tp} + \text{FP}) \quad (18)$$



(a)



(b)

Fig. 5 Fake images generated by regularized bad generators where (a) 500 and (b) 1000 labeled training images were used for SVHN classifiers

4.3 Experiment settings

The implemented method was executed on a laptop equipped with an Intel Core i7 12700H processor boasting a clock speed of 4.00 GHz, 32 GB of RAM, and an 8 GB NVIDIA GeForce RTX 3070 graphics card. Python programming language, along with the Torch library, was leveraged to implement the proposed algorithm. Our approach entailed a multi-class classifier, a bad generator, and a discriminator, integrating individual-relational consistency regularization

and a pseudo-labeling technique. To ensure a fair comparison with the referenced paper [5], we adjusted the multi-class classifier by employing a convolutional network with 13 layers (CNN-13) tailored for the SVHN, CINIC-10, COVID-19, and CIFAR-10 datasets. The architectural designs of the bad generator and discriminator networks were inspired by the infoGAN approach [30]. Detailed visual representations of the three component architectures are depicted in Fig. 3.

The images utilized in the networks for the SVHN, CINIC-10, COVID-19, and CIFAR-10 datasets were of dimensions 32×32 . In these datasets, batch sizes and the length of the z vector were set to 100, as stipulated in reference [5]. Concerning learning rates, the classifier, discriminator, and generator were assigned values of 0.1, $2e-4$, and $2e-4$, respectively, across all datasets. For the SVHN dataset, a variance (σ) value of 0.03 was employed in the local deviation function, whereas for the CIFAR-10 and CINIC-10 datasets, variance values of 0.07 and 1 were chosen, respectively. These specific variance values were meticulously selected to optimize performance on each respective dataset. The training regimen for all datasets comprised 400 epochs. For more detailed information regarding the model's parameters, kindly consult Table 2.

4.4 Results

Experiments were conducted on the SVHN, CINIC-10, COVID-19, and CIFAR-10 datasets. The process of selecting images for training the classifier actively involved choosing 500 and 1000 images from the SVHN training data as labeled images, as detailed in reference [5]. In Fig. 4, confusion matrices for CNN-13 classifiers on the SVHN test data are illustrated. When trained with 500 labeled images, the classifier demonstrated TP values of 1698, 4786, 4039, 2772, 2467, 2280, 1931, 1918, 1578, and 1528 for the ten output classes on the SVHN test data. Nevertheless, with an increase in labeled images to 1000 during the training process, the classifier exhibited improved TP values for the same classes, achieving 1693, 4859, 4060, 2750, 2464, 2289, 1930, 1920, 1576, and 1531.

Figure 5 displays a series of bad fake images created by regularized-bad generators. The classifiers on SVHN were trained using 500 and 1000 labeled training images, as illustrated in (a) and (b), respectively. Incorporating both individual and relational consistency techniques significantly improves the performance of the bad generator. This enhancement results in the generation of informative bad fake images, offering valuable insights into the classifier's decision boundary. These techniques enable the bad generator to produce fake images that effectively depict and convey information about the classifier's decision boundaries.

Fig. 6 Confusion matrices for classifiers trained with (a) 1000 and (b) 4000 labeled training images on CIFAR-10 test data

Actual Label	0	1	2	3	4	5	6	7	8	9
0	894	7	39	5	1	3	16	4	18	13
1	3	961	0	1	0	3	2	0	5	25
2	32	0	964	26	118	70	51	6	1	2
3	21	0	19	686	22	183	44	13	5	7
4	7	0	11	10	894	36	25	16	1	0
5	7	0	18	72	20	854	9	20	0	0
6	6	0	10	10	4	5	961	3	1	0
7	12	1	10	14	14	34	9	903	1	2
8	34	7	21	1	1	1	4	2	920	9
9	6	27	1	3	0	4	6	1	4	948
	0	1	2	3	4	5	6	7	8	9

(a)

Actual Label	0	1	2	3	4	5	6	7	8	9
0	963	1	7	3	0	2	3	3	16	2
1	2	970	0	0	0	1	0	0	5	22
2	37	0	886	18	14	19	21	4	1	0
3	8	1	8	860	9	77	22	6	6	3
4	1	0	3	12	961	6	8	9	0	0
5	2	0	13	60	19	886	4	14	1	1
6	4	0	5	8	2	3	975	3	0	0
7	3	0	5	16	11	7	0	956	1	1
8	19	6	1	2	0	0	0	0	967	5
9	11	20	0	0	0	1	2	0	8	958
	0	1	2	3	4	5	6	7	8	9

(b)

In our experiment, we curated a subset of labeled images from the CIFAR-10 training dataset, specifically selecting 1000 and 4000 images to train the classifier [5]. Figure 6 illustrates the confusion matrices for classifiers on the CIFAR-10 test data. Upon training the CNN-13 classifier with 1000 labeled images, we observed TP values for the ten output classes on the CIFAR-10 test data, yielding corresponding TP values of 894, 961, 694, 686, 894, 854, 961, 903, 920, and 948. Conversely, training the classifier with 4000 labeled images resulted in improved TP values for the same classes. The enhanced TP results were 963, 970, 886, 860, 961, 886, 975, 956, 967, and 958. Moreover, improvements were noted in FP, TN, and FN values. Figure 7 illustrates the generated fake images by regularized-bad generators using 1000 and 4000 labeled training images for classifiers on CIFAR-10, as depicted in subfigures a and b, respectively. The application of individual-relational consistency regularization to the bad images significantly contributes to enhancing the diversity of the generated images.

In another experiment conducted on the CINIC-10 dataset, we selected 7,000 and 10,000 training images for classifier training, while the remaining data were designated as unlabeled images. Figure 8 illustrates the confusion matrices

on the CINIC-10 test data for the CNN-13 classifier trained with 7,000 and 10,000 labeled images. Initially, the TP values for the ten output classes of the test data were estimated as 7893, 7424, 7200, 7111, 7028, 5599, 8192, 7640, 7920, and 7385 using a classifier trained with 700 labels per class. However, these TP values improved to 7819, 7795, 7559, 7001, 7126, 6807, 8445, 8089, 8078, and 7351 using a classifier trained with 1,000 labels per class. Figure 9 depicts the bad images generated by individual-relational consistent bad generators on the CINIC-10 dataset. In this case, the generators engage in an adversarial game with the classifiers, which were trained using 700 and 1000 labeled images per class, respectively.

From the COVID-19 training data, we selected 10 and 40 labeled images per class for training the classifiers, while the remaining data was utilized as unlabeled. The confusion matrices for the CNN-13 classifiers on the COVID-19 test data are depicted in Fig. 10. The classifier trained with 30 labeled images achieved TP values of 34, 36, and 32, while the classifier trained with 120 labeled images estimated TP values of 35, 36, and 35. Figure 11 shows the bad images generated using individual-relational consistent bad generators on the COVID-19 dataset.



Fig. 7 Fake images generated by regularized bad generators where (a) 1000 and (b) 4000 labeled training images were used for CIFAR-10 classifiers, respectively

The proposed consistency regularization losses introduce increased variation and diversity in the generated bad images, resulting in a broader spectrum of informative images. These bad generated images display inter-class features, incorporating characteristics or elements from multiple classes. Consequently, the classifier exhibits low confidence in accurately classifying these images due to the presence of conflicting or ambiguous features. They offer valuable support to the classifier in precisely delineating the decision boundary and enhancing the pseudo-labeling

of low-confidence unlabeled images. Additionally, Fig. 12 illustrates the accuracy curves depicting the convergence of the model on the SVHN, CINIC-10, CIFAR-10, and COVID-19 datasets. These curves visually represent the model's accuracy progression during training.

5 Discussion

In this section, we offer a comprehensive discussion of our findings, commencing with a quantitative analysis in Sect. 5.1, followed by the outcomes of our ablation studies in Sect. 5.2, and concluding with a qualitative discussion in Sect. 5.3.

5.1 Quantitative discussion

The error rates obtained using the CNN-13 classifier for the CIFAR-10 and SVHN datasets are summarized in Table 3. Non-bold results in Tables 3 and 4, drawn from papers [5, 8, 39], emphasize accuracy and error rate metrics. Conversely, the bold numerical values in the tables denote the best results attained from our implementation under identical conditions.

On the SVHN dataset, the state-of-the-art (SOTA) supervised model (Manifold Mixup) yielded estimated error values of 20.57 ± 0.63 and 13.07 ± 0.53 when trained with 500 and 1000 labeled images, respectively [67]. Nonetheless, when the same labeled and unlabeled images were utilized in the semi-supervised method (ICT) with the same classifier, substantially lower error rates of 4.23 ± 0.15 and 3.89 ± 0.04 were achieved [5]. In contrast, the SNTG method [23] yielded error estimates of 3.99 ± 0.24 and 3.86 ± 0.27 . Furthermore, when our proposed IRC-BSGAN method was applied to the SVHN dataset, a notable improvement in the mean error rate was observed, with achieved values of 3.87 ± 0.17 and 3.67 ± 0.09 , respectively. Regarding other performance metrics for SVHN test images, using 500 labeled images resulted in a precision of 96.08%, a sensitivity of 96.02%, a specificity of 99.51%, and an F1 score of 96.03%. When the number of labeled images increased to 1000, a precision of 96.33%, a sensitivity of 96.31%, a specificity of 99.54%, and an F1 score of 96.32% were achieved.

Estimated error values of 34.58 ± 0.37 and 18.59 ± 0.18 on the CIFAR-10 dataset were achieved by the SOTA-supervised model (Manifold Mixup) when trained with 1000 and 4000 labeled images, respectively [67]. Under the same conditions, error rates of 15.48 ± 0.78 and 7.29 ± 0.02 were yielded by the semi-supervised ICT method [5]. The CoUA approach [42] predicted error values as 13.83 ± 0.51 and 8.01 ± 0.28 . Furthermore, our proposed IRC-BSGAN method

Fig. 8 Confusion matrices on the CINIC-10 test images via classifiers trained with (a) 700 and (b) 1000 labeled images per class

0	7893	74	115	34	31	24	17	115	563	134
1	71	7427	39	28	11	55	22	70	172	1105
2	456	15	7200	402	225	210	251	97	113	31
3	90	16	214	7111	256	745	251	186	69	62
4	78	11	224	606	7028	380	146	398	95	34
5	75	32	252	1585	707	5599	115	475	105	55
6	45	6	291	280	51	69	8192	21	41	4
7	71	31	105	249	369	332	26	7640	79	98
8	319	102	109	72	66	73	40	116	7920	183
9	142	1087	35	41	23	45	14	55	173	7385
	0	1	2	3	4	5	6	7	8	9

(a)

0	7819	124	130	232	21	23	21	85	466	79
1	21	7795	22	17	6	25	16	38	114	946
2	243	16	7559	271	179	256	298	51	104	23
3	32	61	221	7001	166	1109	226	80	49	55
4	42	37	199	405	7126	517	139	430	78	27
5	36	81	205	943	390	6807	98	322	74	44
6	20	18	207	150	33	89	8445	13	22	3
7	34	47	68	92	220	313	14	8089	55	68
8	212	144	122	76	60	56	38	91	8078	123
9	68	1245	21	28	16	39	13	77	142	7351
	0	1	2	3	4	5	6	7	8	9

(b)

further improved the predicted error rates to 12.76 ± 0.31 and 6.17 ± 0.05 . For CIFAR-10, when 1000 labeled images were utilized, test images exhibited an estimated precision of 87.43%, a sensitivity of 87.15%, a specificity of 98.57%, and an F1 score of 87.03%. As the number of labeled images increased to 4000, the achieved results improved, demonstrating a precision of 93.83%, a sensitivity of 93.82%, a specificity of 99.31%, and an F1 score of 93.80%.

Table 4 provides a comprehensive evaluation of performance results on the CINIC-10 dataset. The outcomes obtained from various approaches, including the proposed method, a basic conventional CNN-13 supervised method, and state-of-the-art semi-supervised learning algorithms [5, 8, 39], are thoroughly analyzed. When the CNN-13 classifier was trained with 700 and 1000 labeled images per class as a supervised model, it resulted in estimated error values of 33.7 ± 0.14 and 30.04 ± 0.29 , respectively. The semi-supervised ICT approach [5], which incorporated unlabeled images, achieved error rates of 25.81 ± 0.16 and 23.19 ± 0.21 . The DSSLDDR+MT model [39] demonstrated error rates of 23.96 ± 0.42 and 21.81 ± 0.16 , while the DNLL model [8] yielded error rates of 22.11 ± 0.28 and

19.38 ± 0.17 . Notably, the proposed IRC-BSGAN method significantly improved the predicted error rates, attaining values of 18.41 ± 0.11 and 15.45 ± 0.13 when used in conjunction with the CNN-13 classifier and 7000 and 10,000 labels, respectively. In the case of CINIC10, with 700 labeled images per class, an estimated precision of 81.62%, a sensitivity of 81.55%, a specificity of 97.95%, and an F1 score of 81.48% were obtained. By increasing the number of labeled images per class to 1000, the achieved results improved, yielding a precision of 84.61%, a sensitivity of 84.52%, a specificity of 98.28%, and an F1 score of 84.52%.

5.2 Ablation studies

In this section, we introduce additional experimental findings aimed at validating the efficacy of the proposed consistency module within the framework of bad generative semi-supervised classification. Furthermore, we investigate how different factors, including the components of the consistency module, the architecture of the classifier, and the choices regarding the activation number of the discriminator, influence the behavior of the system.



Fig. 9 Bad images generated by consistent bad generators a and b on the CINIC-10 dataset where classifiers were trained with (a) 7000 and (b) 10,000 labeled images, respectively

Table 5 presents the average percentages of accuracy obtained from five runs for the COVID-19 test dataset, focusing on the activation number and the corresponding

Fig. 10 Confusion matrices on the COVID-19 test images via classifiers trained with (a) 30 and (b) 120 labeled images

Actual Label	0	34	0	2
	1	0	36	0
	2	0	4	32
		0	1	2
		Predicted Label		

(a)

Actual Label	0	35	0	1
	1	0	36	0
	2	0	1	35
		0	1	2
		Predicted Label		

(b)

number of labeled images (30, 120). The table displays four different activation numbers: 500, 100, 700, and 300. Each activation number is associated with two sets of accuracy percentages, one for each label. For the activation number of 500, the average accuracy percentages for the two labels (30, 120) are $92.66\% \pm 0.17\%$ and $97.53\% \pm 0.53\%$, respectively. When the activation number is set to 100, the accuracy percentages reduce to values of $92.28\% \pm 0.54\%$ and $97.41\% \pm 0.46\%$ for the two labels. Similarly, an activation number of 300 yields higher accuracy percentages of $94.45\% \pm 0.23\%$ and $98.26\% \pm 0.31\%$ for the respective labels. Lastly, an activation number of 700 results in accuracy percentages of $93.98\% \pm 0.27\%$ and $97.71\% \pm 0.29\%$ for the two labels. These findings suggest that the choice of activation number has an impact on the accuracy of the COVID-19 test dataset. Three hundred activation numbers generally lead to improved accuracy, as demonstrated by the increasing percentages observed in the table.

Table 6 presents the average percentages of accuracy obtained from five runs for the COVID-19 test dataset, focusing on consistency modules and different classifiers. The table depicts two models, bad GAN and IRC-BSGAN, and two classifiers, CNN-13 and WRN-28-2. The accuracy percentages are provided for two scenarios: COVID-19, with 30 and 120 labeled images. For the base model bad GAN and classifier CNN-13, the average accuracy percentages for COVID-19 with 30 labels are $92.59\% \pm 0.36\%$, and with 120 labels, they are $97.18\% \pm 0.15\%$. On the other hand, when using the IRC-BSGAN with the same classifier CNN-13, the accuracy percentages increase to $94.45\% \pm 0.23\%$ for COVID-19 with 30 labels and $98.26\% \pm 0.31\%$ with 120 labels.

Similarly, for the basic model bad GAN and classifier WRN-28-2 with 30 labels, the accuracy percentages are $92.13\% \pm 0.66\%$, and with 120 labels, they are $96.76\% \pm 0.65\%$. Utilizing the consistency module with the classifier WRN-28-2, the accuracy percentages reach $93.51\% \pm 0.93\%$ with 30 labels and $97.25\% \pm 0.43\%$ with 120 labels. These results indicate that the choice of consistency module and classifier has a notable impact on the accuracy of the COVID-19 test dataset. The bad GAN with the individual-relation-consistency module (IRC-BSGAN) generally outperforms the basic bad GAN module, resulting in higher

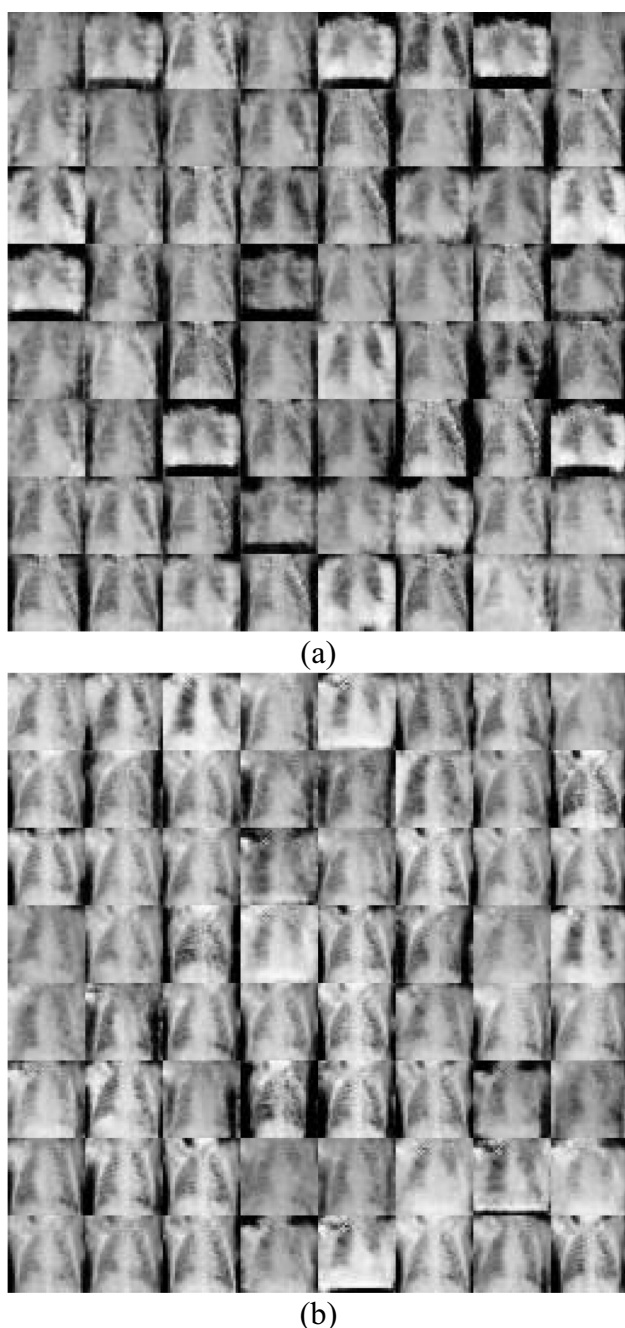


Fig. 11 Bad images generated by consistent bad generators a and b on the COVID-19 dataset where classifiers were trained with (a) 30 and (b) 120 labeled images, respectively

accuracy percentages. Additionally, the classifier CNN-13 tends to demonstrate better accuracy compared to WRN-28–2 in most cases. However, it's important to note that these findings are based on an average of five runs and may vary depending on the specific dataset and experimental setup.

Table 7 provides the performance metrics for the IRC-BSGAN method on the COVID-19 test dataset using

classifiers trained with (a) 30 labeled images and (b) 120 labeled images. The average performance across all classes using 30 labeled images is a precision of 94.71%, a sensitivity of 94.44%, a specificity of 97.22%, and an F1 score of 94.44%. On the other hand, when utilizing 120 labeled images, the average performance across all classes is a precision of 98.15%, a sensitivity of 98.15%, a specificity of 99.07%, and an F1 score of 98.14%. These results confirm that increasing the number of labeled images from 30 to 120 leads to improved performance across all metrics.

Based on Table 8, the activation number that yields the lowest average error rate is 500 for both the SVHN and CIFAR-10 datasets. For the SVHN dataset, this activation number results in an error rate of $3.67\% \pm 0.09\%$. Similarly, for the CIFAR-10 dataset, the error rate decreases to $12.76\% \pm 0.31\%$ when using the same activation number.

5.3 Qualitative discussion

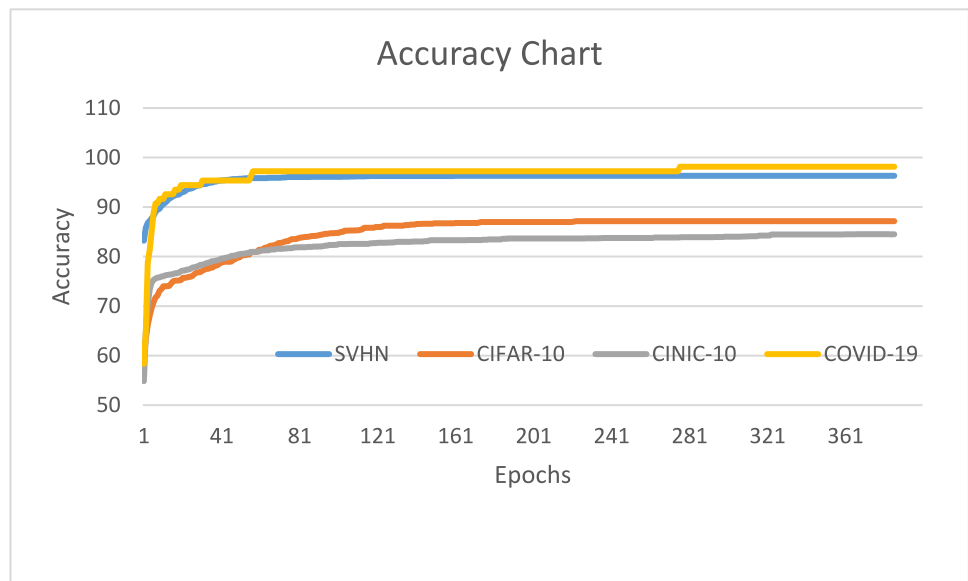
The results of the study demonstrate a significant enhancement in classifier performance when unlabeled data is integrated using our proposed semi-supervised method (IRC-BSGAN). Moreover, an increase in the number of labeled images employed for model training correlates with a reduction in the error rate. This finding highlights the importance of access to a larger pool of labeled images in enhancing the model's capacity for generalization and accurate predictions. The theoretical implications of our approach hold particular relevance in the context of improving bad GANs' performance. By incorporating unlabeled images, including those of low quality, and leveraging techniques such as data augmentation and individual-relational consistency regularization, we can elevate the overall effectiveness of these models. This highlights the significance of integrating diverse data sources and enforcing consistency in model predictions, thereby enhancing their robustness and reliability.

The practical implications of our method extend to various domains, notably medical diagnosis, where obtaining labeled data can be arduous due to the requirement for expert annotations [12]. Additionally, the method is relevant for applications like fraud detection or anomaly detection, where accurately determining the decision boundary is paramount for effectively identifying fraudulent or anomalous instances [69].

Our proposed method, referred to as individual-relational consistency for bad semi-supervised generative adversarial networks (IRC-BSGAN), distinguishes itself from previous works in several key aspects:

1. Novel latent (inversed) individual and relational consistency: Previous methods have primarily focused on individual consistency regularization, which perturbs individual data points to improve classification performance. In contrast, the suggested method introduces

Fig. 12 The accuracy curves on the SVHN, CINIC-10, CIFAR-10 and COVID-19 data



(inversed) individual and relational consistency regularization losses applied to bad fake images to enhance their generation and detect boundaries for low-confidence unlabeled images. Relational consistency regularization operates on the latent vectors of bad fake images in low-density areas to enhance coherence and consistency within the latent space and the generated images. This approach improves feature learning in mislabeled images near the decision boundary, reducing incorrect pseudo-labeling using insights from bad fake images.

2. Leveraging low-density areas: Our proposed method specifically targets low-density areas where the model encounters difficulties in capturing meaningful information and generating bad images. By promoting local smoothing of each fake image and encouraging smoothness among neighboring points (relational), the method aims to uncover additional semantic information from unlabeled images, especially low-confidence unlabeled images in low-density regions.

Table 3 Average percentages of error rates obtained from five runs for the SVHN and CIFAR-10 test datasets

Reference	Model	Number Of labels			
		SVHN (500)	SVHN (1000)	CIFAR-10 (1000)	CIFAR-10 (4000)
[67]	Supervised (Manifold Mixup)	20.57 ± 0.63	13.07 ± 0.53	34.58 ± 0.37	18.59 ± 0.18
[35]	π model	6.65 ± 0.53	4.82 ± 0.17	31.65 ± 1.20	12.36 ± 0.31
[35]	TempEns	5.12 ± 0.13	4.42 ± 0.16	23.31 ± 1.01	12.16 ± 0.24
[37]	MT	4.18 ± 0.27	3.95 ± 0.19	21.55 ± 1.48	12.31 ± 0.28
[45]	VAT	-	5.42 ± NA	-	11.36 ± NA
[45]	VAT + Ent	-	3.86 ± NA	-	10.55 ± NA
[46]	VAdD	-	4.16 ± 0.08	-	11.32 ± 0.11
[23]	SNTG	3.99 ± 0.24	3.86 ± 0.27	18.41 ± 0.52	10.93 ± 0.14
[37]	MT + Fast SWA	-	-	15.58 ± NA	9.05 ± NA
[39]	DSSLDDR	9.07 ± 1.15	6.65 ± 0.27	27.42 ± 1.46	15.98 ± 0.68
[30]	bad GAN	-	4.25 ± NA	20.63 ± NA	14.41 ± NA
[39]	DSSLDDR + MT	4.12 ± 0.47	3.9 ± 0.28	16.99 ± 0.57	10.41 ± 0.2
[43]	Triple-BigGAN	-	-	-	8.90 ± 0.21
[42]	CoUA	-	-	13.83 ± 0.51	8.01 ± 0.28
[40]	DAGE	-	-	14.00 ± 0.49	9.01 ± 0.24
[41]	ReliaMatch	-	7.13 ± 0.28	-	7.42 ± 0.05
[5]	ICT	4.23 ± 0.15	3.89 ± 0.04	15.48 ± 0.78	7.29 ± 0.02
Current research	IRC-BSGAN	3.87 ± 0.17	3.67 ± 0.09	12.76 ± 0.31	6.17 ± 0.05

Table 4 Average error rate percentages from five runs for the CINIC-10 test data

Reference	Model	Number Of labels	
		CINIC-10 (700)	CINIC-10 (1000)
[68]	Supervised (CNN-13)	33.7 ± 0.14	30.04 ± 0.29
[35]	π model	29.66 ± 1.12	27.04 ± 0.85
[35]	TE	30.38 ± 1.01	27.35 ± 0.86
[37]	MT	28.41 ± 0.29	25.71 ± 0.12
[5]	ICT	25.81 ± 0.16	23.19 ± 0.21
[39]	DSSLDDR	29.35 ± 0.31	26.75 ± 0.24
[39]	DSSLDDR + MT	23.96 ± 0.42	21.81 ± 0.16
[8]	DNLL	22.11 ± 0.28	19.38 ± 0.17
Current research	IRC-BSGAN	18.41 ± 0.11	15.45 ± 0.13

Table 5 Average percentages of accuracy obtained from five runs for the COVID-19 test dataset depending on the activation number

The activation number	Number Of labels	
	30	120
500	92.66 ± 0.17	97.53 ± 0.53
100	92.28 ± 0.54	97.41 ± 0.46
300	94.45 ± 0.23	98.26 ± 0.31
700	93.98 ± 0.27	97.71 ± 0.29

Table 6 Average percentages of accuracy obtained from five runs for the COVID-19 test dataset depending on consistency modules and different classifiers

Model	classifier	Number Of labels	
		30	120
bad GAN	CNN-13	92.59 ± 0.36	97.18 ± 0.15
IRC-BSGAN	CNN-13	94.45 ± 0.23	98.26 ± 0.31
bad GAN	WRN-28-2	92.13 ± 0.66	96.76 ± 0.65
IRC-BSGAN	WRN-28-2	93.51 ± 0.93	97.25 ± 0.43

Table 7 The performance metrics for the IRC-BSGAN method on the COVID-19 test dataset via classifiers trained with (a) 30 and (b) 120 labeled images

(a)				
Class number	Precision	Sensitivity	Specificity	F1
COVID-19	100	94.44	100	97.14
pneumonia	90.00	100	94.44	94.74
health	94.12	88.89	97.22	91.43
Average	94.71	94.44	97.22	94.44
(b)				
Class number	Precision	Sensitivity	Specificity	F1
COVID-19	97.22	97.22	98.61	97.22
pneumonia	100	100	100	100
health	97.22	97.22	98.61	97.22
Average	98.15	98.15	99.07	98.14

Table 8 Average error rate (%) obtained from five runs for the SVHN and CIFAR-10 test datasets depending on the activation number

The activation number	Data set	
	SVHN	CIFAR-10
500	3.67 ± 0.09	12.76 ± 0.31
100	5.39 ± 0.21	15.71 ± 0.46
300	5.15 ± 0.11	15.43 ± 0.20
700	4.69 ± 0.15	14.36 ± 0.52

6 Conclusion

Semi-supervised learning has emerged as a promising approach to leverage both labeled and unlabeled images for classification tasks. However, existing methods often assign pseudo-labels based solely on high probabilities, neglecting the relationships among low-confidence images. In response, this study proposes an innovative method called individual-relational consistency for bad semi-supervised generative adversarial networks (IRC-BSGAN) to address the challenges of incorrect pseudo-labeling in semi-supervised learning scenarios. By incorporating novel individual and relational consistency regularization techniques, the proposed method aims to enhance the generation of bad fake images that resemble support vectors and improve the detection of the decision boundary.

The contributions of this research include the introduction of (inversed) individual and relational consistency regularization losses on bad fake images and their latent vectors, as well as leveraging low-density areas to uncover additional semantic information. Experimental results conducted on the SVHN, CINIC-10, COVID-19, and CIFAR-10 datasets demonstrated the effectiveness of IRC-BSGAN, outperforming existing approaches in terms of classification performance. IRC-BSGAN provided not only valuable insights but also delivered promising results for tackling the challenges of semi-supervised classification, thereby opening up avenues for further research and advancements in this field.

While our research yields promising results, it is crucial to acknowledge potential limitations or challenges. These may include scalability issues when applying our method to larger datasets, sensitivity to hyperparameters, increased computational demands, or difficulties in generalizing the approach to diverse domains. It is prudent to consider these aspects and potential limitations when implementing and evaluating IRC-BSGAN in practical applications.

Author contributions Mohammad Saber Iraj: Writing – original draft- review & editing, Methodology, Investigation, Validation, Conceptualization. Jafar Tanha: Supervision, Conceptualization, Methodology, Writing – review & editing, validation. Mohammad-Ali Balafar: Supervision, Validation. Mohammad-Reza Feizi-Derakhshi: Supervision, Validation.

Data availability Data available on request.

Declarations

Ethical and informed consent This article does not contain any studies with human participants or animals performed by any of the authors. The datasets used in the manuscript are derived from publicly available data sets and may be obtained from the appropriate authors upon reasonable request.

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Khan MA et al (2021) A deep survey on supervised learning based human detection and activity classification methods. *Multimed Tools Appl* 80(18):27867–27923
- Duarte JM et al (2021) Deep analysis of word sense disambiguation via semi-supervised learning and neural word representations. *Inf Sci* 570:278–297
- Gao Y et al (2020) A semi-supervised convolutional neural network-based method for steel surface defect recognition. *Robot Comput-Integr Manuf* 61:101825
- Zhang B et al (2021) Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling. *Adv Neural Inf Process Syst* 34:18408–18419
- Verma V et al (2022) Interpolation consistency training for semi-supervised learning. *Neural Netw* 145:90–106
- Kang P et al (2022) Intra-class low-rank regularization for supervised and semi-supervised cross-modal retrieval. *Appl Intell* 52(1):33–54
- Han M et al (2023) A survey of multi-label classification based on supervised and semi-supervised learning. *Int J Mach Learn Cybern* 14(3):697–724
- Xu H et al (2023) Semi-supervised learning with pseudo-negative labels for image classification. *Knowl-Based Syst* 260:110166
- Li S et al (2023) Robust Teacher: Self-correcting pseudo-label-guided semi-supervised learning for object detection. *Comput Vis Image Underst* 235:103788
- Peng Z et al (2023) Semi-supervised medical image classification with adaptive threshold pseudo-labeling and unreliable sample contrastive loss. *Biomed Signal Process Control* 79:104142
- Li D, Liu Y, Song L (2022) Adaptive weighted losses with distribution approximation for efficient consistency-based semi-supervised learning. *IEEE Trans Circuits Syst Video Technol* 32(11):7832–7842
- Wang X et al (2021) Deep virtual adversarial self-training with consistency regularization for semi-supervised medical image classification. *Med Image Anal* 70:102010
- Shi Y et al (2023) Multi-granularity knowledge distillation and prototype consistency regularization for class-incremental learning. *Neural Netw* 164:617–630
- Su L et al (2023) Dual consistency semi-supervised nuclei detection via global regularization and local adversarial learning. *Neurocomputing* 529:204–213
- Poon H-K et al (2019) Hierarchical gated recurrent neural network with adversarial and virtual adversarial training on text classification. *Neural Netw* 119:299–312
- Chen Y et al (2022) Generating robust real-time object detector with uncertainty via virtual adversarial training. *Int J Mach Learn Cybern* 13(2):431–445
- Yan M, Hui SC, Li N (2023) DML-PL: Deep metric learning based pseudo-labeling framework for class imbalanced semi-supervised learning. *Inf Sci* 626:641–657
- Ke B, Lu H, You C, Zhu W, Xie L, Yao Y (2024) A semi-supervised medical image classification method based on combined pseudo-labeling and distance metric consistency. *Multimedia Tools and Applications* 83(11):33313–33331
- Feng Z et al (2022) Dmt: Dynamic mutual training for semi-supervised learning. *Pattern Recogn* 130:108777
- Bi X et al (2022) Entropy-weighted reconstruction adversary and curriculum pseudo labeling for domain adaptation in semantic segmentation. *Neurocomputing* 506:277–289
- Duan Y et al (2020) Mutexmatch: semi-supervised learning with mutex-based consistency regularization. *IEEE Trans Neural Netw Learn Syst* 35:8441–8455
- Liu Q et al (2020) Semi-supervised medical image classification with relation-driven self-ensembling model. *IEEE Trans Med Imaging* 39(11):3429–3440
- Luo Y, Zhu J, Li M, Ren Y, Zhang Bl (2018) Smooth neighbors on teacher graphs for semi-supervised learning. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 8896–8905
- Van Engelen JE, Hoos HH (2020) A survey on semi-supervised learning. *Mach Learn* 109(2):373–440
- Li X, Luan Y, Chen L (2023) Semi-supervised GAN with similarity constraint for mode diversity. *Appl Intell* 53(4):3933–3946
- Wang L, Sun Y, Wang Z (2022) CCS-GAN: a semi-supervised generative adversarial network for image classification. *Vis Comput* 38(6):2009–2021
- Mayer C, Paul M, Timofte R (2021) Adversarial feature distribution alignment for semi-supervised learning. *Comput Vis Image Underst* 202:103109
- Li W, Wang Z, Li J, Polson J, Speier W, Arnold CW (2019) Semi-supervised learning based on generative adversarial network: a comparison between good GAN and bad GAN approach. In: *CVPR Workshops*, pp 1–11
- Dai Z, Yang Z, Yang F, Cohen WW, Salakhutdinov RR (2017) Good semi-supervised learning that requires a bad gan. *Adv Neural Inf Process Syst* 30:4–6
- Dong J, Lin T (2019) MarginGAN: adversarial training in semi-supervised learning. *Adv Neural Inf Process Syst* 32:2–5
- Zhang Y et al (2021) Twin self-supervision based semi-supervised learning (TS-SSL): Retinal anomaly classification in SD-OCT images. *Neurocomputing* 462:491–505

32. Gong Y, Wu Q, Cheng D (2023) A co-training method based on parameter-free and single-step unlabeled data selection strategy with natural neighbors. *Int J Mach Learn Cybern* 14(8):2887–2902
33. Tian Y et al (2022) Consistency regularization teacher–student semi-supervised learning method for target recognition in SAR images. *Vis Comput* 38(12):4179–4192
34. Chen J, Yang M, Ling J (2021) Attention-based label consistency for semi-supervised deep learning based image classification. *Neurocomputing* 453:731–741
35. Zhou W et al (2020) Mutual improvement between temporal ensembling and virtual adversarial training. *Neural Process Lett* 51:1111–1124
36. Ding W, Abdel-Basset M, Hawash H (2021) RCTE: A reliable and consistent temporal-ensembling framework for semi-supervised segmentation of COVID-19 lesions. *Inf Sci* 578:559–573
37. Feng C et al (2020) Domain adaptation with SBADA-GAN and Mean Teacher. *Neurocomputing* 396:577–586
38. Liu L, Tan RT (2021) Certainty driven consistency loss on multi-teacher networks for semi-supervised learning. *Pattern Recogn* 120:108140
39. Yang M et al (2023) Discriminative semi-supervised learning via deep and dictionary representation for image classification. *Pattern Recogn* 140:109521
40. Tu E et al (2022) Deep semi-supervised learning via dynamic anchor graph embedding in latent space. *Neural Netw* 146:350–360
41. Jiang T et al (2023) Reliamatch: Semi-supervised classification with reliable match. *Appl Sci* 13(15):8856
42. Huo X et al (2023) Collaborative learning with unreliability adaptation for semi-supervised image classification. *Pattern Recogn* 133:109032
43. Gangwar A et al (2023) Triple-BigGAN: Semi-supervised generative adversarial networks for image synthesis and classification on sexual facial expression recognition. *Neurocomputing* 528:200–216
44. Chen L et al (2020) Seqvat: Virtual adversarial training for semi-supervised sequence labeling. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics
45. Miyato T et al (2018) Virtual adversarial training: a regularization method for supervised and semi-supervised learning. *IEEE Trans Pattern Anal Mach Intell* 41(8):1979–1993
46. Park S, Park J, Shin SJ, Moon IC (2018) Adversarial dropout for supervised and semi-supervised learning. In: *Proceedings of the AAAI conference on artificial intelligence* 32(1)
47. Sohn K et al (2020) Fixmatch: Simplifying semi-supervised learning with consistency and confidence. *Adv Neural Inf Process Syst* 33:596–608
48. Yamaguchi T, Murakawa M (2022) Mixup gamblers+: Learning interpolated pseudo “uncertainty” in latent feature space for reliable inference. *Pattern Recogn Lett* 164:191–199
49. Jahanyar B, Tabatabaee H, Rowhanimanesh A (2023) MS-ACGAN: A modified auxiliary classifier generative adversarial network for schizophrenia’s samples augmentation based on microarray gene expression data. *Comput Biol Med* 162:107024
50. Struski Ł et al (2022) Locogan—locally convolutional gan. *Comput Vis Image Underst* 221:103462
51. Toutouh J et al (2023) Semi-supervised generative adversarial networks with spatial coevolution for enhanced image generation and classification. *Appl Soft Comput* 148:110890
52. Contreras-Cruz MA et al (2023) Generative Adversarial Networks for anomaly detection in aerial images. *Comput Electr Eng* 106:108470
53. Huang C et al (2023) A review of deep learning in dentistry. *Neurocomputing* 554:126629
54. Zhang Y et al (2023) Integrated intelligent fault diagnosis approach of offshore wind turbine bearing based on information stream fusion and semi-supervised learning. *Expert Syst Appl* 232:120854
55. Peng D et al (2021) SAM-GAN: Self-Attention supporting Multi-stage Generative Adversarial Networks for text-to-image synthesis. *Neural Netw* 138:57–67
56. Wu X et al (2022) Face aging with pixel-level alignment GAN. *Appl Intell* 52(13):14665–14678
57. Xu W, Jiang L, Li C (2021) Improving data and model quality in crowdsourcing using cross-entropy-based noise correction. *Inf Sci* 546:803–814
58. Xie Y, Zhang J, Xia Y (2019) Semi-supervised adversarial model for benign–malignant lung nodule classification on chest CT. *Med Image Anal* 57:237–248
59. Garg S et al (2020) A unified view of label shift estimation. *Adv Neural Inf Process Syst* 33:3290–3300
60. Huo X, Zhang Y, Wu S (2024) Semi-supervised class-conditional image synthesis with Semantics-guided Adaptive Feature Transforms. *Pattern Recogn* 146:110022
61. Qi Z et al (2021) Pccm-gan: Photographic text-to-image generation with pyramid contrastive consistency model. *Neurocomputing* 449:330–341
62. Netzer Y, Wang T, Coates A, Bissacco A, Wu B, Ng AY (2011) Reading digits in natural images with unsupervised feature learning. *NIPS workshop on deep learning and unsupervised feature learning* 2011(2):4
63. Canayaz M (2021) MH-COVIDNet: Diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images. *Biomed Signal Process Control* 64:102257
64. Krizhevsky A, Hinton G (2009) Learning multiple layers of features from tiny images
65. Xu J, Zhang Y, Miao D (2020) Three-way confusion matrix for classification: A measure driven view. *Inf Sci* 507:772–794
66. Deepak S, Ameer P (2019) Brain tumor classification using deep CNN features via transfer learning. *Comput Biol Med* 111:103345
67. Verma V et al (2019) Manifold mixup: Better representations by interpolating hidden states. in *International conference on machine learning*. PMLR
68. Chen J, Yang M, Gao G (2020) Semi-supervised dual-branch network for image classification. *Knowl-Based Syst* 197:105837
69. Xia X et al (2022) GAN-based anomaly detection: A review. *Neurocomputing* 493:497–535

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



Mohammad Saber Iraj He is a faculty member of the Computer Science Department at PNU University in Iran. His research areas include semi-supervised classification, pattern recognition, image processing, feature selection, fuzzy logic, and game theory.



Mohammad Ali Balafar He is currently a professor at the Faculty of Computer Engineering, University of Tabriz, Iran. His research interests include feature extraction, biomedical MRI, brain, convolutional neural nets, face recognition, fuzzy set theory, image classification, image enhancement, image motion analysis, image representation, image resolution, image segmentation, image sensors, image sequences, image texture, medical image processing, Bayes methods, probability, wavelet transforms.



Jafar Tanha He obtained his Bachelor's degree in Computer science in December 1999 and his Master's degree in Computer Science and Applied Mathematics in June 2002 from the Department of Computer Science and Mathematics of the University of Amir Kabir (Polytechnic Tehran). He received a Ph.D. degree in the Computer Science Department from the University of Amsterdam. He is currently a professor at the Faculty of Computer Engineering, University of Tabriz, Iran. His research interests include learning (artificial intelligence), pattern classification, semi-supervised learning, pattern clustering, complex networks, and game theory.



Mohammad-Reza Feizi-Derakhshi Mohammad-Reza Feizi-Derakhshi received a B.S. degree in software engineering from the University of Isfahan, Iran, and an M.Sc. and Ph.D. degree in artificial intelligence from the Iran University of Science and Technology, Tehran, Iran. He is currently a professor at the Faculty of Computer Engineering, University of Tabriz, Iran. His research interests include natural language processing, optimization algorithms, deep learning, social network analysis, and intelligent databases.