

Received January 18, 2018, accepted February 13, 2018, date of publication February 23, 2018, date of current version March 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2808938

cC-GAN: A Robust Transfer-Learning Framework for HEp-2 Specimen Image Segmentation

YUEXIANG LI^{ID} AND LINLIN SHEN, (Member, IEEE)

College of Computer Science and Software Engineering, Computer Vision Institute, Shenzhen University, Shenzhen 518060, China

Corresponding author: Linlin Shen (llshen@szu.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61672357, Grant 61702339, and Grant U1713214, in part by the Science Foundation of Shenzhen under Grant JCYJ20160422144110140, and in part by the China Postdoctoral Science Foundation under Grant 2017M622779.

ABSTRACT Human epithelial type 2 (HEp-2) cell images play an important role for the detection of antinuclear autoantibodies in autoimmune diseases. As the HEp-2 cell has hundreds of different patterns, none of currently available HEp-2 datasets contain all of the types. Therefore, existing automatic processing systems for HEp-2 cells, e.g., cell segmentation and classification, needs to be transferred between different data sets. However, the performances of transferred system often dramatically decrease, especially when transferring supervised-approaches, e.g., deep learning network, from large dataset to the small but similar ones. In this paper, a novel transfer-learning framework using generative adversarial networks (cC-GAN) is proposed for robust segmentation of different HEp-2 datasets. The proposed cC-GAN tries to solve the overfitting problem of most deep learning networks and improves their transfer-capacity. An improved U-net, so-called Residual U-net (RU-net), is developed to work as the generator for cC-GAN model. The cC-GAN was first trained and tested using I3A dataset and then directly evaluated using MIVIA dataset, which is much smaller than I3A. The segmentation result demonstrates the excellent transferring-capacity of our cC-GAN framework, i.e., a new state-of-the-art segmentation accuracy of 75.27% was achieved on MIVIA without finetuning.

INDEX TERMS Cell segmentation, generative adversarial networks, fully convolutional network.

I. INTRODUCTION

The patterns of Human Epithelial type 2 (HEp-2) cell provide useful information for the diagnosis of autoimmune diseases. To alleviate the shortcoming of traditional manual approaches, e.g. inter-observer variability, various automatic systems for HEp-2 cell classification have been proposed [1]. Foggia gave a thorough review of automatic processing systems developed for HEp-2 cell images in [2] and [3] and more recent reviews were provided by Hobson *et al.* [4]–[6]. As most of the existing classification approaches involve cell segmentation as their first step, the accuracy of segmentation makes direct influence to the final classification accuracy. Hence, a robust automatic segmentation approach for HEp-2 cell image is worthwhile to develop.

Cell image segmentation is a field that has been extensively studied. The accurate segmentation results can increase the performance of subsequent processing, e.g. cell classification, etc. Intensity thresholding is the oldest, and still one of the most predominant approaches for cell

segmentation [7]. Perner proposed the first thresholding based segmentation for HEp-2 cell images using Otsu [8]. Based on the previous work, Huang developed a hybrid segmentation method combining Otsu and Canny operator to improve the segmentation accuracy [9]. In more recent research, a novel approach for HEp-2 cell segmentation based on the framework of verification-based multi-threshold probing was proposed by Jiang *et al.* [10]. Other typical segmentation approaches, e.g. morphological filtering, have also been applied to HEp-2 cell images. For example, Percannella employed a sort of classifier-controlled morphological dilation operation for the better structuring of HEp-2 cell areas [11]. Many researches [12]–[14] improved the typical watershed algorithm [15] for the segmentation of HEp-2 cell images. In more recent research, Merone *et al.* proposed an active-contour based segmentation algorithm for HEp-2 images [16]. Although many studies [16], [17] have been made for HEp-2 cell segmentation, due to the big variances of appearances among different HEp-2 cell categories, most

of the previous studies are only able to provide accurate segmentation for images of some specific cell patterns.

Deep convolutional neural networks (CNNs) gain extensive attention from researchers since their performances surpassed the state-of-the-art in many visual recognition tasks. Although the typical application of convolutional networks is classification, many researchers have made their trials using CNN to address the problem of biomedical image segmentation in recent years. Ciresan et al. trained a CNN in a sliding-window setup to predict the class label, i.e. foreground/background, of each pixel [18]. The network produces outstanding segmentation results for Electron Microscopy (EM) images. However, there is a trade-off between the window-size and segmentation performance of the CNN approach, i.e. larger window-size decreases the segmentation accuracy whilst smaller size provide limited context. Hence, inspired by the Fully Convolutional Network (FCN) [19], Ronneberger proposed the U-net for microscopy image segmentation [20], which surpassed the CNN-based segmentation network on EM dataset. In more recent research, Li proposed a hybrid network combining the FCN and Residual Network to segment HEp-2 specimen image [21], which achieved a new state-of-the-art segmentation accuracy. The main idea under FCN model is to apply the classification networks (AlexNet [22], VGG net [23], GoogLeNet [24], and ResNet [25]) to segmentation task by transforming the last classifier layers to deconvolutional layers. Compared to the CNN-based segmentation network, FCN is a more elegant architecture as it does not require the setting of sliding-window size [20]. Due to the effective segmentation results produced by FCN, the model has been used to segment not only microscopy images but also other medical images. For example, Chen proposed a FCN-based network for the segmentation of glands [26]. Recently, generative adversarial networks (GAN) has also been applied to generate segmentation mask in [27]. Kohl employed the architecture proposed in [27] to segment the MRI of aggressive prostate cancer [28].

Although the deep learning approaches outperformed the traditional methods on various processing tasks, the primary problems of using deep learning to process HEp-2 cell image gradually emerge, i.e. lack of training data and network overfitting. As the HEp-2 cell has hundreds of different patterns, the image volume of some patterns may be relatively small, e.g. only four Cytoplasmic specimen images are available in MIVIA dataset [2]. Insufficient training data makes the deep network unable to correctly segment and recognize the cell images of rare patterns. To address the problem, models are often pre-trained on large dataset, e.g. ImageNet or I3A datasets [4], and then transferred to the small but similar ones, e.g. MIVIA. Phan finetuned a model pre-trained on ImageNet to MIVIA dataset for HEp-2 cell classification [29]. Bayramoglu et al. pre-trained their network on I3A to achieve better classification performance on MIVIA [30]. However, due to the overfitting problem, the performances of transferred deep learning system often dramatically decrease.

Furthermore, most of previous transfer-learning work mainly focuses on the classification problem of small HEp-2 dataset. Annotated masks were used to extract cell areas from specimen images in the classification approaches. Therefore, their classification performances may significantly decreases in practical application due to the lack of reliable automatic segmentation approach. The first work transferring deep learning network to segment small HEp-2 dataset was made in the very recent research [21]. A fully convolutional residual network (FCRN) pre-trained on I3A dataset was transferred to segment MIVIA dataset, with a segmentation accuracy of 59.88%. Although the transferred FCRN surpassed all the traditional methods, the robustness of transferred HEp-2 segmentation network still has large margin for improvement.

In this paper, we proposed a novel model using conditional generative adversarial networks (cC-GAN) for robust segmentation among different HEp-2 datasets. The networks trained with proposed cC-GAN framework alleviate the influence caused by network overfitting and achieved the better transferring capacity for small-scale HEp-2 dataset. A hybrid network combining ResNet and U-net (RU-net) is proposed as the generator of cC-GAN. The RU-net achieved better segmentation performances than the typical U-net on the HEp-2 datasets of I3A and MIVIA. To more effectively train the transfer-learning frameworks, a novel hybrid training scheme (additional-epoch training, AEt) is proposed, which consists of common training and extra GAN training. As the GAN model involves two primary components, i.e. Generator network (G) and Discriminator network (D), we also investigated the relationship between the G and D, and disclosed the rules of network selections for the small-scale HEp-2 dataset, e.g. MIVIA.

II. RELATED WORK

A. GENERATIVE ADVERSARIAL NETWORKS

Generative adversarial networks (GAN) is first proposed by Goodfellow [31] in 2014, which provides a game-theoretic formulation for training models for image generation. The GAN attracts increasing attentions from researchers in recent years. Extensive works improving the typical GAN model has been proposed to increase the quality and variability of generated images [32], [33].

A generative adversarial network (GAN) consists of two networks trained in the oppositions. The typical generator (G) takes a random noise vector z and accordingly generates an image, $X_{fake} = G(z)$. The typical discriminator (D) receives an input image and decides if it is a training image (y) or synthesized image from generator ($G(z)$). The objective of a typical GAN can be expressed as

$$\begin{aligned} L_{GAN}(G, D) = & E_{y \sim p_{data}(y)}[\log D(y)] \\ & + E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \end{aligned} \quad (1)$$

where G tries to minimize the objective whilst the discriminator (D) tries to maximize it, i.e. $\text{argmin}_G \text{max}_D L_{GAN}(G, D)$.

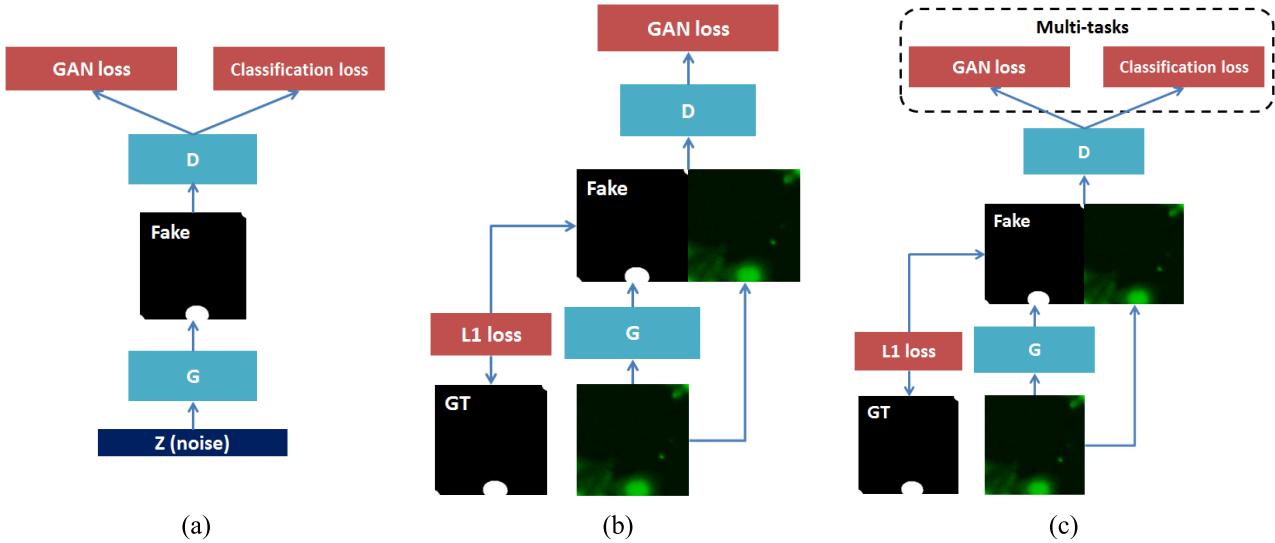


FIGURE 1. Comparison of GAN frameworks. G and D stand for generator and discriminator, respectively. (a) The typical AC-GAN (GAN with auxiliary classifier). The classification loss is used to improve the quality of generated images. (b) The typical pix2pix model with two losses, i.e. GAN loss and L_1 loss. (c) The proposed cC-GAN with GAN loss, L_1 loss and classification loss.

B. GAN WITH AUXILIARY CLASSIFIER (AC-GAN)

Recent research [34] illustrates that the performance of image synthesis can be improved by adding an auxiliary classifier to generative adversarial networks. Based on the observation, Odena proposed a variant of GAN, namely AC-GAN [34]. The AC-GAN added a classification branch to the discriminator of GAN. Therefore, the discriminator has both a probability distribution over sources and a probability distribution over the class labels. The objective function of AC-GAN has two parts, i.e. L_{GAN} and L_{Class} . The L_{Class} can be generally defined as

$$L_{Class} = E_{y \sim p_{class}(y)} [\log P(C = c | X_{real})] + E_{z \sim p_z(z)} [P(C = c | X_{fake})] \quad (2)$$

where D is trained to maximize $L_{GAN} + L_{Class}$ while G is trained to maximize $L_{Class} - L_{GAN}$.

C. IMAGE-TO-IMAGE TRANSLATION USING GAN

In addition to image generation, image-to-image translation using GAN models is another on-going research topic. It applies GAN in the conditional setting [35], [36] to perform various kinds of image transformations, e.g. colorizing pictures with object edges, transferring day photos to night and converting maps to aerial photos, etc. The first work in this area was named by *pix2pix* [27]. Assuming the provided conditional image as x , the objective of conditional GAN (cGAN) involved in *pix2pix* can be expressed as

$$L_{cGAN}(G, D) = E_{x,y \sim p_{data}(x,y)}[\log D(x, y)] + E_{x \sim p_{data}(x), z \sim p_z(z)}[\log(1 - D(x, G(x, z)))] \quad (3)$$

The pix2pix framework also employs a L_1 distance loss between the synthesized image and corresponding ground

truth for generator network, i.e. $L_{L1}(G)$. Hence, the final objective of pix2pix framework is $\operatorname{argmin}_{G} \max_D L_{cGAN}(G, D) + \lambda L_{L1}(G)$. λ is a ratio factor for L_1 loss (L_{L1}).

III. NETWORK ARCHITECTURE

A. THE FRAMEWORK OF cC-GAN

We proposed novel conditional generative adversarial networks with classifier (cC-GAN) to improve the transferring capacity of pre-trained fully-convolutional network. Fig. 1 (c) presents the architecture of cC-GAN, which contains two primary components, i.e. generator (G) and discriminator (D). Given the input cell image, the generator in our cC-GAN model accordingly generates the cell masks, which is equivalent to cell segmentation whilst the discriminator aims to distinguish the generated images from ground truths.

The proposed cC-GAN is a hybrid network based on AC-GAN (Fig. 1 (a)) [34] and pix2pix model (Fig. 1 (b)) [27]. The typical AC-GAN is proposed to synthesize an image from a random noise. Thus, it is not able to complete the image-to-image translation task, e.g. from cell images to binary masks. The pix2pix model is able to perform image-to-image translation task. However, it only uses the GAN loss to assess the translation performance, which is not stable and easily collapses in training process. Since the AC-GAN demonstrates that classification loss can improve the quality of samples generated by GAN model and stabilize GAN training, our cC-GAN (Fig. 1 (c)) combines AC-GAN and pix2pix models by adding a classification branch to the discriminator of pix2pix.

Three losses are involved in the proposed cC-GAN, i.e. L_1 loss (L_{L1}), GAN loss (L_{cGAN}) and Softmax loss (L_{class}). The L_1 loss is employed to encourage generator network to predict the right label, i.e. foreground/background for each pixel in the image. The GAN loss is used to discriminate the output

TABLE 1. Architecture of discriminator in cC-GAN.

Layer	Type	Kernel size, number & stride
1	I-C- LeakyReLU	4x4, 64, 2
2	C-BN-LeakyReLU	4x4, 128, 2
3	C-BN-LeakyReLU	4x4, 256, 2
4	C-BN-LeakyReLU	4x4, 512, 2
5	C-Sigmoid	4x4, 1, 1
5	FC-Softmax	7

of generator network from the ground-truth. With the GAN loss, the discriminator network is able to tell the generator whether the entire generated image or a large portion of it mismatches with the ground-truth. Therefore, the generator can accordingly update its network weights. The GAN loss can be seen as a regularization term for the generator network to avoid overfitting. The Softmax loss, defined in Eq. (4), enables the discriminator network to predict the categories of input cell images.

$$L_{class} = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_i -\log\left(\frac{e^{f_i}}{\sum_j e^{f_j}}\right) \quad (4)$$

where f_j denotes the j -th element ($j \in [1, K]$, K is the number of classes) of vector f predicted by the discriminator, l_i is the label of i -th input feature and N is the number of training data.

Hence, according to Eq. (2-3), the objective of our cC-GAN framework can be expressed as

$$\begin{aligned} \arg(\min_G \max_D (L_{cGAN}(G, D) + \lambda L_1(G)) \\ + \min_G \min_D L_{class}(G, D)) \end{aligned} \quad (5)$$

where λ is set to 100 in the experiments.

The proposed cC-GAN adopts the architectures of U-net/RU-net as generator network, which will be introduced in the next section. The architecture for discriminator is similar to the one introduced in [27], whose detail is listed in Table 1. The pipeline consists of Input layer (I), Convolutional layer (C), Batch Normalization layer (BN), LeakyReLU, Fully-connected layer (FC) and loss layers, i.e. Sigmoid and Softmax.

B. RESIDUAL U-NET BASED GENERATOR

U-net is a fully-convolutional network widely-used as generator for conditional GAN models, e.g. pix2pix, which fuses the features extracted from shallow and deep convolutional layers to produce accurate segmentation masks. Residual Network (ResNet) won the competition of ILSVRC 2015, demonstrating its excellent capacity of feature extraction. Combining the advantages of the two network architectures, we proposed so-called Residual U-net (RU-net) for the generator network of our cC-GAN. The architecture of proposed RU-net is presented in Fig. 2.

The proposed RU-net consists of 7 residual modules (RM) and 7 deconvolutional layers (DC). The purple rectangles are the output feature maps from corresponding residual modules. The black arrows stand for concatenation operation. The

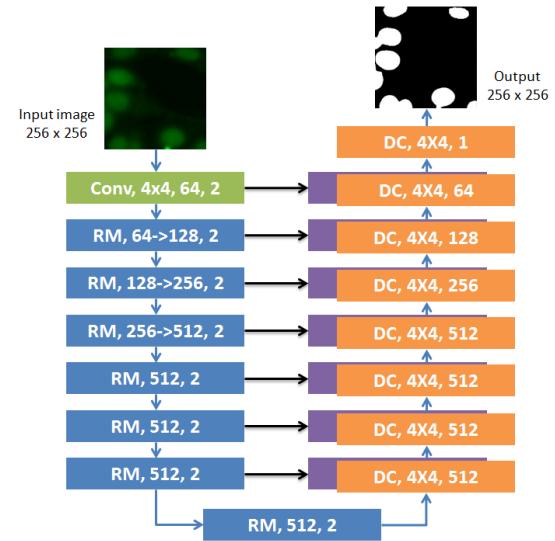


FIGURE 2. Flowchart of RU-net. Each rectangle represents a network module which may contain more than one layer. The numbers in each module represent the size and amount of convolutional kernels. The steps of convolutional layer and residual modules are set to 2.



FIGURE 3. The architecture of residual module.

convolutional layer and residual modules in RU-net extract image features by making convolutional operation with a stride of 2. To produce the segmentation results of the same size to input image, the deconvolutional layers are employed to up-sampled the feature maps. Batch normalization [37] is used to reduce internal covariate shift during training process. LeakyReLU [38] is used as the non-linear activation function for the proposed RU-net.

As deeper network has larger receptive field enclosing more contextual information for segmentation, the proposed RU-net replaces the original plain convolutional connections of U-net with residual modules. The RM structure was verified to help alleviating the gradient vanishing problem as the network goes deeper. As shown in Fig. 3, the skip connections can directly propagate signals forward and backward.

C. ADDITIONAL-EPOCH TRAINING SCHEME (AE-TRAINING)

When the training scheme stated in [27] and [28] simultaneously trains the G and D, we found it not stable to train the cC-GAN in our experiments. Therefore, we proposed in this paper the additional-epoch training scheme (AEt) to train our cC-GAN model:

Step 1: The generator, i.e. U-net/RU-net, (without GAN structure) was first trained on I3A dataset only supervised by L1 loss until it converges, i.e. 9 epochs in our experiments.

Step 2: The pre-trained generator was loaded into the cC-GAN model and further finetuned on I3A dataset for 1 to 2 epochs with the supervision of GAN, L1 and classification losses.

The aim of GAN model is to balance the performances between generator and discriminator. Hence, the GAN model is easy to collapse when one component (G/D) always outperforms the other. The proposed AE-training scheme separately trains the generator and discriminator, such that the oscillations occurred in network training caused by random initialization can be reduced. That helps to balance the generator and discriminator, stabilize the training process, alleviate the model collapse problem and improve the segmentation performance of generator.

D. IMPLEMENTATION

The proposed framework is developed with *Torch* toolbox. The networks were trained on one Tesla K80 (11GB RAM). The detailed setting for training the cC-GAN model is accordingly set as claimed in [27]. ‘Adam’ [39] is used as the solver for Stochastic Gradient Descent.

IV. PROCESSING APPROACHES

A. PRE-PROCESSING

The I3A dataset contains 1008 HEp-2 specimen images with size 1388×1040 and their contrast is low. An appropriate image pre-processing approach is necessary for the generator network to better extract internal features and perform segmentation. Hence, the same pre-processing approach stated in [21] is adopted in our work.

Since the proposed cC-GAN model can be trained with few training data, our data augmentation scheme only involves image cropping. We consecutively cropped 20 pieces of 256×256 sub-images from each HEp-2 specimen image and generated a set of 20,160 images.

B. POST-PROCESSING (OPTIONAL)

As most of the generated segmentation results have overlapped cells as shown in Fig. 4 (a), the proposed cC-GAN framework employs the watershed transform to separate them. The positions of seed points for watershed transform

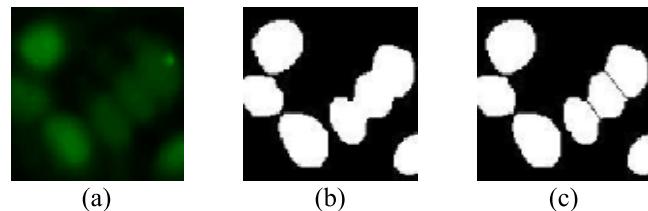


FIGURE 4. Disjoint connected cells. (a) Patch of original cell image. (b) Initial segmentation results from cC-GAN model. (c) Separation results from watershed transform.

are detected by applying morphological erosion to the initial segmentation result (Fig. 4 (b)). Fig. 4 (c) is the final result separating the connected three cells. The post-processing step is optional.

V. EXPERIMENTAL RESULTS

A. DATASET

1) I3A¹

The dataset was first released in the fluorescent image based cell classification contest organized by ICIP 2013 [4], and used again in the contest organized by ICPR 2014 [5]. The dataset records 252 specimens from seven categories, Homogeneous (53), Speckled (52), Nucleolar (50), Centromere (51), Golgi (10), Nuclear membrane (21), and Mitotic spindle (15). The numbers in brackets are the number of specimens for corresponding type of cells. For each specimen, four images were captured in different locations with size of 1388×1040 ; therefore, I3A HEp-2 dataset contains 1008 grayscale specimen images.

2) MIVIA²

MIVIA dataset, first used in ICPR 2012 Contest, is widely-accepted for the performance evaluation of HEp-2 cell segmentation algorithm. The dataset consists of 28 HEp-2 green-channel specimen images with a resolution of 1388×1038 , which can be classified into six categories: Homogeneous (6), Fine speckled (4), Coarse speckled (5), Centromere (6), Nucleolar (4) and Cytoplasmic (4).

It is worthwhile to mention that MIVIA dataset only contains 28 specimen images which are too few to finetune a deep learning framework. Furthermore, the cell categories are different between I3A and MIVIA, resulting in difficulties for a deep learning network pre-trained on I3A to yield accurate segmentation results for images of cell category only available in MIVIA, e.g. Cytoplasmic.

B. EVALUATION CRITERION

The cell areas were manually annotated as ground truths by the dataset provider. Using the provided ground truths, the segmentation accuracy (SEG) can be calculated via:

$$SEG = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN} \quad (7)$$

where the values of TP (number of true positives), FP (number of false positives) and FN (number of false negatives) were computed according to the definitions proposed in previous work [11].

C. EVALUATION OF cC-GAN

We first use I3A dataset (five-fold cross validation) to elaborate the training process of cC-GAN. The split protocol of

¹<http://nerone.diem.unisa.it/hep2-benchmarking/dbtools/>

²<http://mivia.unisa.it/>

five-fold cross validation is the same to [21]. The U-net and RU-net are employed as generator network. The proposed AE-training mechanism, i.e. 9 epochs of common training + 1 epoch of GAN training, was adopted as the training scheme. The typical cGAN only using GAN loss [31] was involved for comparison.

The curves of training error during extra GAN epoch for one of the five-fold cross validation are presented in Fig. 5. The print frequency for training losses is set to 50 iterations; therefore, we have 350 records. The errG and errD stand for the losses from generator and discriminator networks, which are $L_{cGAN}(G, D)$ for typical cGAN and $L_{cGAN}(G, D) + L_{class}(G, D)$ for our cC-GAN. As illustrated in Fig. 5, compared to the proposed cC-GAN (Fig. 5 (c) and (d)), the performances of generator and discriminator are not balanced during extra cGAN training (Fig. 5 (a) and (b)), i.e. the discriminator always surpasses the generator.

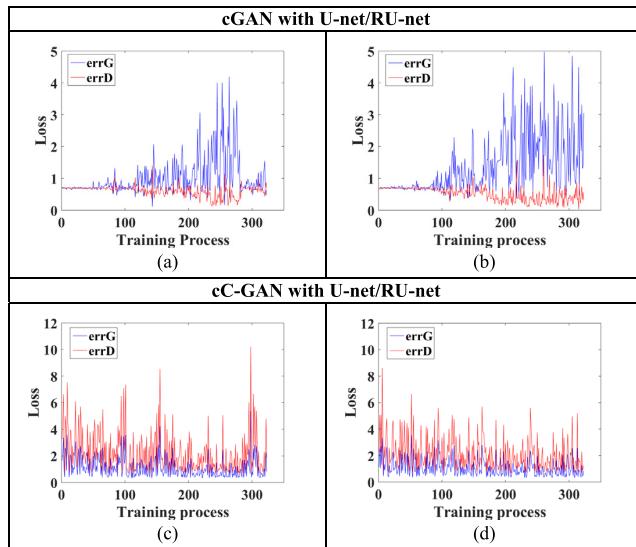


FIGURE 5. Training error curves during extra GAN epoch. (a) cGAN with U-net (b) cGAN with RU-net (c) cC-GAN with U-net (d) cC-GAN with RU-net.

We listed in Table 2 the SEG of five-fold cross validation for different GAN models. One can observe that the performance of U-net and RU-net gradually increases from epoch #3 to #9 as the networks start to converge. They reached the best after 9 epochs, i.e. 85.56% for U-net and 86.00% for RU-net. Due to the unstable training process of cGAN, performance degradations are observed after additional cGAN training, i.e. drops of 2.75% for U-net and 6.28% for RU-net. In contrast, the extra epoch of cC-GAN training provides improvements of SEG, i.e. 0.21% for U-net and 0.15% for RU-net.

To assess the transferring capacity, the pre-trained cC-GAN framework was directly applied to MIVIA dataset without finetuning. Fig. 6 presents the transferring capacities of cGAN and cC-GAN on MIVIA, with different training epochs. The lines in red are the SEG for GAN models using RU-net and the cyan lines are for the ones using U-net.

TABLE 2. Variation of SEG during AE-training on I3A (%).

	Training Epoch			Extra GAN training	
	3	6	9	cGAN	cC-GAN
U-net	84.66	85.27	85.56	82.81	85.77
RU-net	84.68	85.34	86.00	79.72	86.15

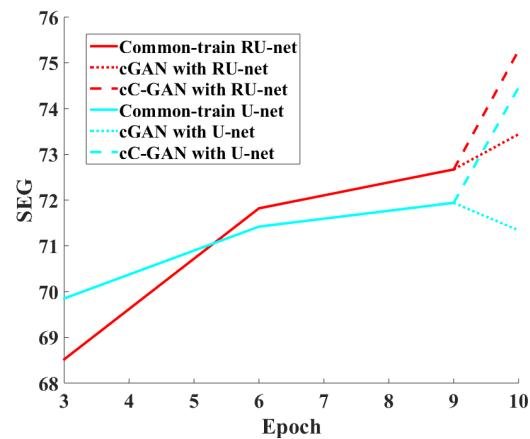


FIGURE 6. SEG of U-net/RU-net on MIVIA dataset during AE-training. The cyan lines stand for GANs with U-net and the red lines are for the ones with RU-net.

Fig. 6 illustrates that RU-net provides better transferring capacity than U-net after 3 epochs of training. After 9 epochs of common training, the SEG of RU-net on MIVIA reaches around 72%, which is about 1% higher than that of U-net. Due to the imbalance AE-training, cGAN produces a marginal improvement for RU-net whilst a degradation for U-net. In comparison, the proposed cC-GAN significantly boosts the performance of U-net/RU-net on MIVIA in epoch #10. The performance improvements are 2.52% and 2.60% for U-net and RU-net, which demonstrates that as the regularization term, the additional epoch of cC-GAN training in AE-training scheme can effectively alleviate the overfitting problem and improve the transferring capacity of U-net/RU-net.

Fig. 7 presents segmentation results of different deep learning frameworks for each cell category of MIVIA. The results from U-net have a severe problem of over-segmentation, e.g. many fault segmentations were generated in the result for Coarse Speckled. The RU-net produces decent results alleviating the problem of over-segmentation, which illustrates that the residual modules can help network to better extract cell features. Compared to the common trained U-net/RU-net, the networks with AE-training, i.e. cC-GAN with U-net/RU-net, produces more complete cell outlines, which can be easily observed by comparing the results of Homogeneous and Centromere. As the cell pattern of Cytoplasmic is not contained in the training set, i.e. I3A, the segmentation results generated by the frameworks for this cell pattern are visually different from the ground truths.

D. COMPARISON WITH STATE-OF-THE-ART

1) I3A DATASET

In this section, we evaluate the segmentation accuracy of proposed cC-GAN with benchmark algorithms on I3A dataset

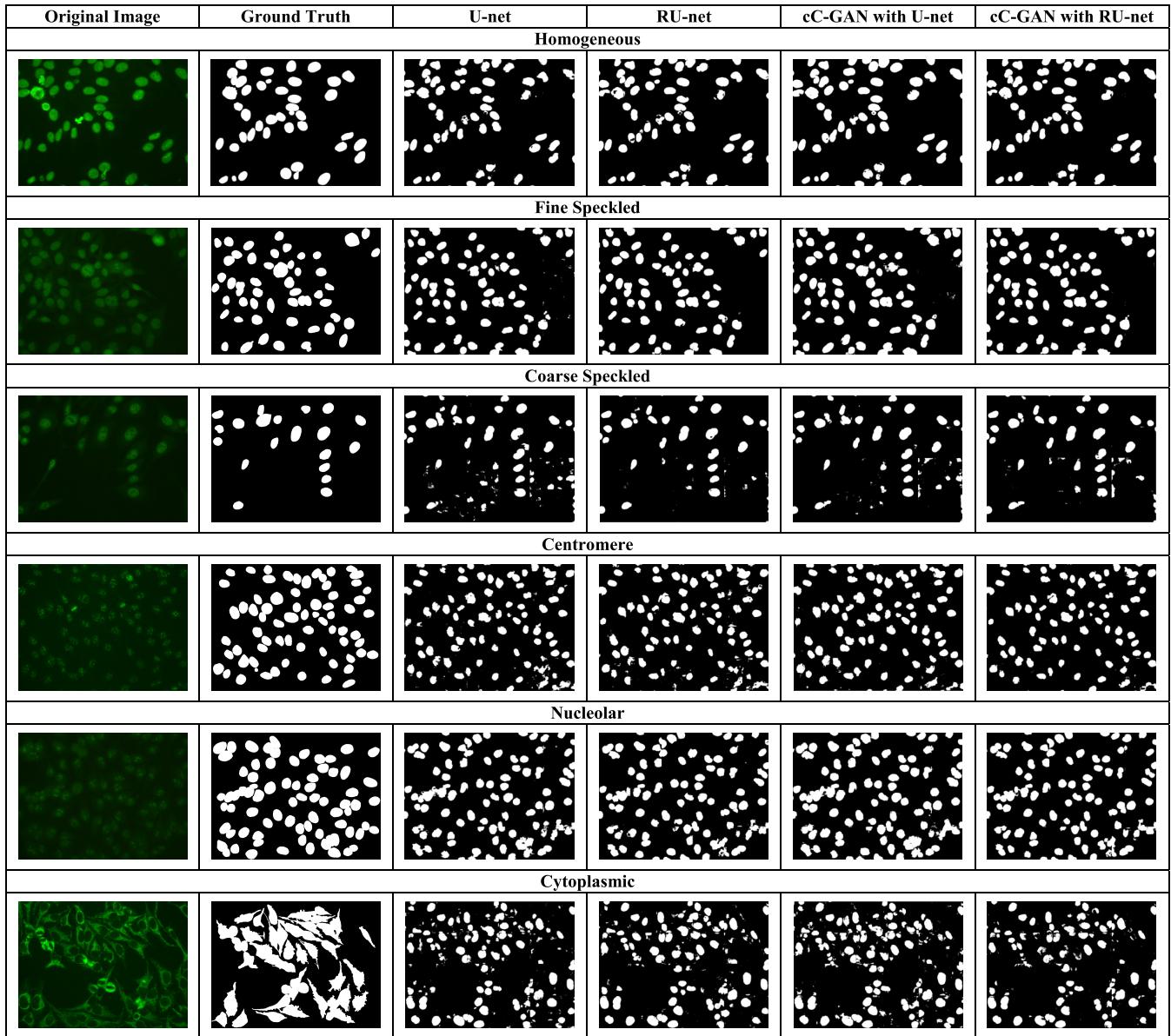


FIGURE 7. Segmentation results on MIVIA dataset. The first column shows examples from each cell category. The corresponding ground truths are presented in the second column. The following columns present the results from different frameworks.

using five-fold cross-validation. The Otsu approach [8], Fully Convolutional ResNet (FCRN) [21], original U-net [20] and original pix2pix model [27] are involved for comparison. Table 3 lists the SEG for different approaches.

It can be observed from Table 3 that all deep learning based approaches perform much better than the Otsu. For non-GAN networks, FCRN-88 with 88 layers performs the best and our RU-net ranks the second. The GAN network proposed for image segmentation, pix2pix, does not perform well, i.e. a SEG of 78.93%. Compared to the adoption of cGAN, which didn't improve the performance of U-net, the proposed cC-GAN improved the performance of both U-net and RU-net.

Due to the larger receptive fields, deeper networks are able to capture more contextual information for complex image

TABLE 3. Segmentation performances on I3A.

	Depth	SEG (%)
Otsu [8]	-	69.98
FCRN-88 [21]	88	87.29
U-net [20]	16	85.56
pix2pix [27]	21	78.93
RU-net (Ours)	23	86.00
Models with AE-training:		
cGAN with U-net (Ours)	21	82.81
cC-GAN with U-net (Ours)	21	85.77
cC-GAN with RU-net (Ours)	28	86.15

processing tasks. Hence, the deepest FCRN-88 produces the best SEG of 87.29%, which is 1.14% higher than that of our cC-GAN with RU-net. However, the deeper architecture

TABLE 4. SEG of different approaches for each category in MIVIA (%).

	Homogeneous	Fine speckled	Coarse speckled	Centromere	Nucleolar	Cytoplasmic	Average SEG
Hand-crafted approaches							
Otsu [8]	76.69	55.86	39.73	49.49	16.13	9.11	41.17
Watershed [12]	49.86	46.73	45.28	56.64	48.12	29.38	46.00
Adaptive method [9]	67.45	61.76	60.16	48.50	49.14	22.99	51.67
Auto-learning [11]	63.61	66.73	68.43	50.11	57.46	36.19	57.09
Marker-control watershed [14]	81.90	82.10	77.40	70.30	71.30	49.30	74.40
Deep-learning approaches							
FCRN-50 [25]	74.82	64.65	64.49	68.32	54.07	26.58	58.82
FCRN-101 [25]	75.68	64.62	62.66	61.09	53.49	25.16	57.12
FCRN-88 [21]	75.90	65.91	65.44	66.98	54.76	30.31	59.88
pix2pix: cGAN with U-net [27]	73.99	66.32	62.95	65.70	62.56	31.70	60.54
U-net [20]	81.24	85.52	82.27	72.61	66.28	43.74	71.94
Residual U-net	82.12	86.03	82.77	73.86	66.06	45.15	72.67
Multi-task RU-net	81.65	86.46	83.61	72.10	65.97	50.03	73.30
Frameworks with AE-training							
cGAN with U-net	78.92	85.59	83.52	71.17	64.25	44.56	71.34
cGAN with RU-net	82.19	86.86	84.45	73.13	67.23	46.79	73.44
cC-GAN with U-net	82.54	87.11	84.52	75.48	68.60	48.52	74.46
cC-GAN with RU-net	82.59	87.09	83.83	77.12	70.93	50.03	75.27

leads to the increase of parameters, which increases the risk of overfitting when the training dataset is insufficient.

2) MIVIA DATASET

The proposed networks trained on I3A are directly tested on MIVIA without finetuning for transferring capacity evaluation. The SEG of different approaches for six cell patterns of MIVIA is calculated and listed in Table 4. Previous methods are included in this paper for comparison: the method using Otsu [8], the watershed-based segmentation method for HEp-2 images [12], the adaptive segmentation approach proposed in [9], the auto-learning method suggested in [11], the marker-control watershed [14], the fully convolutional network transformed from typical ResNet-50/101 [25], the fully convolutional ResNet for HEp-2 images [21], the original pix2pix framework [27] and the typical U-net [20]. The last column in Table 4 presents the average SEG of six cell classes. All of the deep learning frameworks listed in Table 4 are trained on I3A and tested on MIVIA without finetuning.

For comparison convenience, we separate Table 4 to three parts, i.e. traditional hand-crafted approaches, deep-learning based approaches and deep-learning framework with AE-training. To investigate the influence caused by classification branch, multi-task RU-net is developed by adding a classification branch to the original RU-net. It can be observed that cC-GAN provides higher improvement, i.e. 2.60%, for RU-net compared to that of multi-task learning, i.e. 0.63%.

For the deep learning models without AE-training, the multi-task RU-net achieves the best segmentation performance on MIVIA, i.e. 73.30%, which is 0.63% higher than the runner-up, i.e. RU-net. The SEG of Multi-task RU-net is 13.42% higher than that of very deep FCRN-88, which

suffered from overfitting problem. The segmentation performance of cGAN with U-net trained with AE-training scheme, i.e. 71.34%, surpasses the original pix2pix, i.e. 60.54%.

Although the proposed AE-training mechanism provides improvements for both U-net and RU-net, its performance varies with different selections of G and D. Due to the difficulty of maintaining balance in AE-training with cGAN, the cGAN-based model yields a marginal improvement of 0.77% to RU-net and a degradation of 0.60% to U-net. The classification branch stables the AE-training by making the discriminator of cC-GAN with a more complex architecture.

Hence, our cC-GAN produces significant improvements for both of U-net and RU-net, i.e. improvements of 2.52% and 2.60%, respectively. The cC-GAN with RU-net trained by AE-training mechanism achieves a new state-of-the-art result, i.e. 75.27% average SEG on MIVIA. The AE-trained cC-GAN (with U-net/RU-net) achieved the best SEG for Homogeneous (82.59%), Fine Speckled (87.11%), Coarse Speckled (84.52%), Centromere (77.12%) and Cytoplasmic (50.03%), which significantly outperforms that of runner-up, i.e. Marker-control watershed. To assess the statistical significance of our results, we made a T-test validation between cC-GAN with RU-net and Marker-control watershed. The probability of observing the given results generated by T-test is 67.60%.

To further evaluate the performance of our cC-GAN, we compared the cC-GAN with benchmarking algorithms on an extra HEp-2 dataset, namely Indirect ImmunoFluorescence Segmentation dataset (IIFS) [16]. The IIFS dataset contains 24 images from the categories of Homogeneous (**H**), Fine speckled (**FS**), Coarse speckled (**CS**) and Nucleolar (**N**). We follow the same protocol presented in [16] for performance evaluation and the results are listed in Table 5. In the

TABLE 5. SEG of Different Approaches for Each Category in MIVIA & IIFS (%).

	H	FS	CS	N	Average
Auto-learning [11]	59.9	58.8	58.3	55.8	58.2
Marker-control watershed [14]	79.5	77.6	76.8	66.0	75.1
Active contour [16]	86.4	85.4	87.0	78.4	84.4
cC-GAN with RU-net	86.8	89.0	87.4	76.4	84.9

table, the segmentation accuracy for each cell category is an average of the results obtained from MIVIA and IIFS.

The proposed cC-GAN was directly applied to IIFS without finetuning. Table 5 shows that the cC-GAN with RU-net achieves comparable accuracy with state-of-the-art, i.e. Active contour, for H, CS and N. For FS, the accuracy of our cC-GAN i.e. 89.0%, is 3.6% higher than that of the runner-up, i.e. Active contour. The average SEG (84.9%) of our cC-GAN is higher than all other competing algorithms. A T-test validation is made between the results from cC-GAN and Active contour to evaluate the statistical significance. The probability generated by T-test is 86.94%.

VI. DISCUSSION AND CONCLUSION

A. DISCUSSION

We have presented a cC-GAN transferring-learning framework to better segment HEp-2 specimen images among different datasets. Based on the presented results, a discussion is presented in this section.

1) HOW TO SELECT MODEL FOR DIFFERENT TASKS?

Table 3 listed the SEG of deep learning networks on I3A for five-fold cross validation. It demonstrates that deeper network can yield better segmentation results if there are sufficient training data. The SEG on I3A dataset increases from 85.56% to 87.29%, when the network depth goes deeper from 16 to 88 layers. However, the deeper network has higher risk of overfitting, e.g. the transferring performance of FCRN-88 is only 59.88% on MIVIA, i.e. 12.06% lower than that of U-net. Hence, very deep networks work better for the application with enough volume of training data, e.g. using the models trained on I3A training set to segment the specimen images in the unpublished I3A testing set.

As small dataset has insufficient data to train or finetune a deep learning network, segmentation framework has to be pre-trained on large dataset and then transferred to segment the small one. The transferring capacity of framework plays a vital role in this task. As presented in Table 4, the proposed AE-training improves the transferring capacity of U-net/RU-net. The core of AE-training is to balance the performances of generator and discriminator networks, which is easier if the two components have similar complexities of network architecture. Hence, to segment the small HEp-2 dataset, e.g. MIVIA, the AE-training cC-GAN with RU-net, which has excellent transferring-capacity, is more preferable.

2) WHY IS cC-GAN BETTER THAN cGAN?

As presented in Table 4, cGAN model decreases the performance of U-net, but produces marginal improvement,

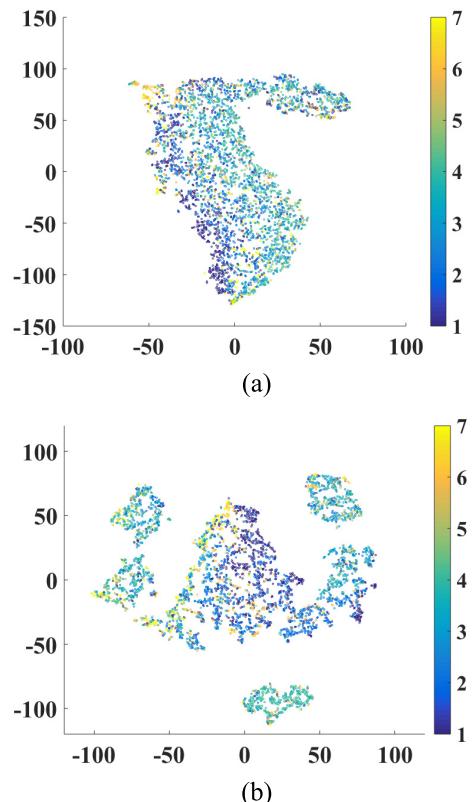


FIGURE 8. The distribution of features learned by different discriminators. (a) cGAN and (b) cC-GAN. Different colors stand for different cell categories.

i.e. 0.77%, for RU-net. In contrast, the proposed cC-GAN provides similar improvements for both U-net and RU-net. The only difference between cGAN and cC-GAN is the extra classification branch added to discriminator. Hence, to further explore the underlying reason of better performances provided by cC-GAN, a discussion of the discriminator networks is performed in this section.

Due to the same structures of first four layers of discriminators in cGAN and cC-GAN, the features learned by layer #4 of discriminator from both GAN frameworks were visualized in Fig. 8 using t-SNE [40]. The numbers in color bar stand for corresponding cell categories, i.e. 1 for Homogeneous, 2 for Speckled, 3 for Nucleolar, 4 for Centromere, 5 for Golgi, 6 for Nuclear membrane and 7 for Mitotic spindle.

Fig. 8 illustrates that the classification branch enriches the information contained in each neuron of discriminator. The features learned by layer #4 of discriminator in cC-GAN (Fig. 8 (b)) are able to not only distinguish real/fake samples, but also recognize their cell categories. As presented in Fig. 8 (b), the cC-GAN has learned to separate Centromere cells from the others, i.e. the cluster of Centromere (label #4) is located in the bottom of Fig. 8 (b). In contrast, the features learned by cGAN have a random distribution (Fig. 8 (a)), which means they have no capacity to distinguish the cell categories. As cC-GAN is only employed for one epoch of

TABLE 6. Leave-one-out on MIVIA (%).

	Homogeneous	Fine speckled	Coarse speckled	Centromere	Nucleolar	Cytoplasmic	Average SEG
cC-GAN with RU-net (without ft)	82.59	87.09	83.83	77.12	70.93	50.03	75.27
cC-GAN with RU-net (with ft)	82.55	89.40	90.16	79.02	73.76	64.43	79.89

training for the purpose of network regularization, its classification performance is not comparable to that of the state-of-the-art algorithms. However, the use of classification branch increases the complexity of discriminator, which stables the AE-training. Furthermore, the classifier branch not only provides richer information for discriminator, but also for the generator, i.e. $\min_G \min_D L_{class}(G, D)$, which enables the generator to produce better segmentation results. Those factors finally result in the improvement of SEG using cC-GAN and the degradation of SEG using cGAN.

3) IS IT POSSIBLE TO USE ULTRA-DEEP NETWORK AS GENERATOR?

We noticed that the ultra-deep network, i.e. FCRN-88 provides excellent performance on I3A dataset with a SEG of 87.29%. According to the performance improvement produced by cC-GAN for U-net/RU-net, there is marginal space for FCRN-88 to improve on MIVIA if the cC-GAN training can be successfully applied. However, in the experiments, we observed that the difficulty of GAN training significantly increases as the generator network goes deeper. Due to the large difference of network complexity between G and D, the model collapse problem, as illustrated in Fig. 5 (b), occurred in GAN training. On the other hand, if we construct a deeper discriminator for FCRN-88, balancing two ultra-deep networks is a difficult task for GAN model. Taking the above factors into consideration, we did not take the ultra-deep network, FCRN-88, as generator for the proposed cC-GAN.

4) FINETUNING OR NOT?

As the MIVIA only has 28 specimen images, which are insufficient for the training of deep learning model, we finetune the model pre-trained on I3A dataset using MIVIA and evaluate the framework performances with leave-one-out experiment. Table 6 listed the results of leave-one-out experiment. Compared to the framework without finetuning (ft), the SEG of Coarse-speckled, Centromere, Nucleolar and Cytoplasmic significantly increased after finetuning. The average SEG of finetuned cC-GAN with RU-net is 79.89%, which is 4.62% higher than that of original framework.

Although the finetuning significantly improves the SEG of several cell patterns, the deep learning networks involved in our work are transferred from I3A to MIVIA without finetuning, which has a number of advantages in the following applications.

While hundreds of HEp-2 cell patterns are available, it is very difficult to collect and label large number of training samples for each pattern. Therefore, models with

knowledge-transferring capacity would be highly useful to directly process small HEp-2 dataset without finetuning. When no training data is provided in practical applications, using pre-trained models to directly process small HEp-2 dataset is also highly desirable.

B. CONCLUSION

Human Epithelial type 2 (HEp-2) cell provides important metric for systemic autoimmune disease diagnosis. To alleviate the influence occurred in traditional manual approaches, a reliable automatic diagnosis system is worthwhile to develop. As the accurate segmentation results can increase the accuracy of sub-sequent processing, e.g. HEp-2 cell classification, the automatic segmentation of HEp-2 images gradually becomes an attracting research topic. However, the variety of HEp-2 cell patterns and the small volume of some patterns make it a difficult task to perform robust segmentation for different HEp-2 datasets.

In this paper, we proposed a novel framework of conditional generative adversarial networks using auxiliary classifier (cC-GAN) to address the challenge of HEp-2 cell segmentation. The cC-GAN networks are employed to alleviate the overfitting problem of most deep learning networks and improve their transfer-capacity. Advanced network architecture, i.e. Residual U-net, is proposed as the generator of our cC-GAN. Furthermore, a novel training scheme, named AE-training, is proposed to further improve the segmentation performance. I3A dataset is used to train the proposed framework and the publicly available much smaller dataset, i.e. MIVIA, is used for testing. A new state-of-the-art segmentation accuracy of 75.27% on MIVIA is achieved by the proposed cC-GAN model without finetuning, which indicates its excellent transferring-capacity between different HEp-2 datasets.

REFERENCES

- [1] L. Shen, J. Lin, S. Wu, and S. Yu, "HEp-2 image classification using intensity order pooling based features and bag of words," *Pattern Recognit.*, vol. 47, no. 7, pp. 2419–2427, Jul. 2014.
- [2] P. Foglia, G. Percannella, P. Soda, and M. Vento, "Benchmarking HEp-2 cells classification methods," *IEEE Trans. Med. Imag.*, vol. 32, no. 10, pp. 1878–1889, Oct. 2013.
- [3] P. Foglia, G. Percannella, A. Saggese, and M. Vento, "Pattern recognition in stained HEp-2 cells: Where are we now?" *Pattern Recognit.*, vol. 47, no. 7, pp. 2305–2314, Jul. 2014.
- [4] P. Hobson, B. C. Lovell, G. Percannella, M. Vento, and A. Willem, "Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset," *Artif. Intell. Med.*, vol. 65, no. 3, pp. 239–250, 2015.
- [5] P. Hobson, B. C. Lovell, G. Percannella, A. Saggese, M. Vento, and A. Willem, "HEp-2 staining pattern recognition at cell and specimen levels: Datasets, algorithms and results," *Pattern Recognit. Lett.*, vol. 82, pp. 12–22, Oct. 2016.

- [6] P. Hobson, B. C. Lovell, G. Percannella, A. Saggese, M. Vento, and A. Wiliem, "Computer aided diagnosis for anti-nuclear antibodies HEp-2 images: Progress and challenges," *Pattern Recognit. Lett.*, vol. 82, pp. 3–11, Oct. 2016.
- [7] E. Meijering, "Cell segmentation: 50 years down the road [life sciences]," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 140–145, Sep. 2012.
- [8] P. Perner, H. Perner, and B. Müller, "Mining knowledge for HEp-2 cell image classification," *Artif. Intell. Med.*, vol. 26, nos. 1–2, pp. 73–161, 2002.
- [9] Y.-L. Huang, Y.-L. Jao, T.-Y. Hsieh, and C.-W. Chung, "Adaptive automatic segmentation of HEp-2 cells in indirect immunofluorescence images," in *Proc. IEEE Int. Conf. Sensor Netw., Ubiquitous Trustworthy Comput.*, Jun. 2008, pp. 418–422.
- [10] X. Jiang, G. Percannella, and M. Vento, "A verification-based multithreshold probing approach to HEp-2 cell segmentation," in *Proc. Comput. Anal. Images Patterns*, 2015, pp. 266–276.
- [11] G. Percannella, P. Soda, and M. Vento, "A classification-based approach to segment HEp-2 cells," in *Proc. Int. Symp. Comput.-Based Med. Syst.*, 2012, pp. 1–5.
- [12] Y.-L. Huang, C.-W. Chung, T.-Y. Hsieh, and Y.-L. Jao, "Outline detection for the HEp-2 cell in indirect immunofluorescence images using watershed segmentation," in *Proc. IEEE Int. Conf. Sensor Netw.*, Jun. 2008, pp. 418–422.
- [13] C.-C. Cheng, J.-S. Taur, T.-Y. Hsieh, and C.-W. Tao, "Segmentation of anti-nuclear antibody images based on the watershed approach," in *Proc. IEEE Conf. Ind. Electron. Appl.*, Jun. 2010, pp. 1695–1700.
- [14] S. Toni, S. D. Cataldo, A. Bottino, and E. Ficarra, "An automated approach to the segmentation of HEp-2 cells for the indirect immunofluorescence ANA test," *Comput. Med. Imag. Graph.*, vol. 40, pp. 62–69, Mar. 2015.
- [15] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.
- [16] M. Merone and P. Soda, "On using active contour to segment HEp-2 cells," in *Proc. IEEE Int. Symp. Comput.-Based Med. Syst.*, Jun. 2016, pp. 118–123.
- [17] C. S. Di, S. Toni, A. Bottino, and E. Ficarra, "ANALyte: A modular image analysis tool for ANA testing with indirect immunofluorescence," *Comput. Methods Programs Biomed.*, vol. 128, p. 86, May 2016.
- [18] D. C. Cirean, G. Alessandro, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 2852–2860.
- [19] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
- [20] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Assist. Intervent.*, 2015, pp. 234–241.
- [21] Y. Li, L. Shen, and S. Yu, "HEp-2 specimen image segmentation and classification using very deep fully convolutional network," *IEEE Trans. Med. Imag.*, vol. 36, no. 7, pp. 1561–1572, Jul. 2017.
- [22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [23] K. Simonyan and A. Zisserman. (2015). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [24] C. Szegedy, W. Liu, Y. Jia, and P. Sermanet, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [26] H. Chen, X. Qi, L. Yu, and P.-A. Heng, "DCAN: Deep contour-aware networks for accurate gland segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2487–2496.
- [27] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros. (2016). "Image-to-image translation with conditional adversarial networks." [Online]. Available: <https://arxiv.org/abs/1611.07004>
- [28] S. Kohl *et al.* (2017). "Adversarial networks for the detection of aggressive prostate cancer." [Online]. Available: <https://arxiv.org/abs/1702.08014>
- [29] H. T. H. Phan, A. Kumar, J. Kim, and D. Feng, "Transfer learning of a convolutional neural network for HEp-2 cell image classification," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Apr. 2016, pp. 1208–1211.
- [30] N. Bayramoglu, J. Kannala, and J. Heikkilä, "Human epithelial type 2 cell classification with convolutional neural networks," in *Proc. IEEE Int. Conf. Bioinf. Bioeng.*, Nov. 2015, pp. 1–6.
- [31] I. J. Goodfellow *et al.*, "Generative adversarial networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 3, 2014, pp. 2672–2680.
- [32] M. Arjovsky, S. Chintala, and L. Bottou. (2017). "Wasserstein GAN." [Online]. Available: <https://arxiv.org/abs/1701.07875>
- [33] D. Berthelot, T. Schumm, and L. Metz. (2017). "BEGAN: Boundary equilibrium generative adversarial networks." [Online]. Available: <https://arxiv.org/abs/1703.10717>
- [34] A. Odena, C. Olah, and J. Shlens. (2016). "Conditional image synthesis with auxiliary classifier GANs." [Online]. Available: <https://arxiv.org/abs/1610.09585>
- [35] M. Mirza and S. Osindero. (2014). "Conditional generative adversarial nets." [Online]. Available: <https://arxiv.org/abs/1411.1784>
- [36] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1060–1069.
- [37] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
- [38] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. Int. Conf. Mach. Learn. Workshop Deep Learn. Audio, Speech, Lang. Process.*, 2013, pp. 1–6.
- [39] D. P. Kingma and J. Ba. (2014). "ADAM: A method for stochastic optimization." [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [40] L. van der Maaten and G. Hinton, "Visualizing high-dimensional data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.



YUXIANG LI received the B.Eng. degree in telecommunications engineering from the Beijing University of Posts and Telecommunications in 2011, the M.S. degree in electronic engineering from The Hong Kong University of Science and Technology in 2012, and the Ph.D. degree in electronic engineering from the University of Nottingham, Nottingham, U.K., in 2016.

He joined the Computer Vision Institute, Shenzhen University, Shenzhen, China, as a Research Fellow, in 2015. His Ph.D. work is in the area of bioelectronics engineering which is an interdisciplinary branch of science requiring knowledge from both bioscience and electronic engineering. The principal aim of his project is to establish a robust monitoring system for cells using image processing techniques related to the area of pattern recognition and computer vision.

Dr. Li attended the Eighth International Symposium on Multispectral Image Processing and Pattern Recognition, Wuhan, China, in 2013, and the conference Optics Within Life Sciences, Ningbo, China, in 2014. He has published two conference papers for these two events.



LINLIN SHEN (M'12) received the Ph.D. degree from the University of Nottingham, Nottingham, U.K., in 2005.

He was a Research Fellow with the Medical School, University of Nottingham, involved in brain image processing of magnetic resonance imaging. He is currently a Professor and a Director of the College of Computer Science and Software Engineering, Computer Vision Institute, Shenzhen University, Shenzhen, China.

His research interests include Gabor wavelets, face/palmprint recognition, medical image processing, and hyperspectral image classification. He was a recipient of the Most Cited Paper Award from the *Journal of Image and Vision Computing* and was the winner of International Competition on Cells Classification by Fluorescent Image Analysis. He is listed as a Most Cited Chinese Researcher by Elsevier.