

引用格式: 卢向日, 汪湛清, 马宏宾. 零和博弈对抗中的代价函数选择与性能评价[J]. 微电子学与计算机, 2021, 38(7): 30-35.
[LU X R, WANG Z Q, MA H B. Cost function selection and performance evaluation in Zero-sum Game confrontation [J]. Microelectronics & Computer, 2021, 38(7): 30-35.]

零和博弈对抗中的代价函数选择与性能评价

卢向日^{1,2}, 汪湛清^{1,2}, 马宏宾^{1,2}

(1 北京理工大学 自动化学院, 北京 100081;

2 复杂系统智能控制与决策国家重点实验室, 北京 100081)

摘 要: 为了解决零和博弈(Zero-sum Game)的 GAN(Generative Adversarial Networks)模型中传统对抗性神经网络的判别器损失值过高问题, 改进了对抗性神经网络的代价函数. 在零和博弈的非完备信息中, 由于博弈对抗的生成器与判别器双方互知的 MNIST 数字图像信息是有限的, 为了恰当的分析出博弈过程中某种因素在对抗中的作用, 代价函数在零和博弈对抗神经网络中显得尤为重要. 对零和博弈对抗环境中的代价函数选择与性能评价进行相关研究, 进而实现获取较大优势的零和博弈. 实验表明: 改进型对抗性神经网络的代价函数优于传统对抗性神经网络损失函数方法.

关键词: 零和博弈; 对抗性神经网络; 非完备信息; 代价函数

中图分类号: TP301.6

文献标识码: A

文章编号: 1000-7180(2021)07-0030-06

Cost function selection and performance evaluation in Zero-sum Game confrontation

LU Xiangri^{1,2}, WANG Zhanqing^{1,2}, MA Hongbin^{1,2}

(1 School of Automation, Beijing Institute of Technology, Beijing 100081, China;

2 State Key Laboratory of Intelligent Control and Decision for Complex Systems, Beijing 100081, China)

Abstract: In order to solve the problem that the discriminator loss value of traditional adversarial neural networks in GAN (Generative Adversarial Networks) model of Zero-sum Game is too high, the cost function of adversarial neural networks is improved. In the incomplete information of the Zero-sum Game, because the MNIST digital image information than what the generator and the discriminator of the game confrontation know each other is limited, to analyze the role of certain factors in the confrontation of the gaming process, the cost function particularly important in the Generative Adversarial Networks model of Zero-sum Game. Research on the selection of cost function and performance evaluation in the confrontation environment of the Zero-sum Game, and then realize the Zero-sum Game with greater advantages. Experimental results show that the cost function of the improved Generative Adversarial Networks is better than the traditional Generative Adversarial Networks loss function method.

Key words: Zero-sum Game; generative adversarial networks; incomplete information; cost function

1 引言

零和博弈(Zero-sum Game), 又称零和游戏, 与

非零和博弈相对, 是博弈论的一个概念, 属非合作博弈. 指参与博弈的各方, 一方的收益必然意味着另一方的损失, 博弈各方的收益和损失相加总和永远为"

零",双方不存在合作的可能. 零和博弈的结果是一方的所得正是另一方的所失,整个社会的利益并不会因此而增加一分.

零和博弈是现在人类社会中严格竞争意义下的模拟. 现在社会竞争的双方较量抽象表现是零和博弈的过程. 现在社会竞争复杂多变,双方对抗越来越依赖人工智能算法分析竞争形势,从而形成客观判断. 机器博弈是人工智能领域重要的研究方向^[1], 2016年Google的Alpha Go战胜韩国围棋大师李世石使得人机博弈变得家喻户晓,接下来升级版的Alpha Go Zero以4:0战胜世界围棋顶尖人物人柯洁等事件揭示了人机博弈的重要意义. 机器博弈根据博弈过程中信息的透明度可以分为完备信息和非完备信息博弈. 完备信息博弈双方信息完全透明,对弈的双方完全了解彼此的博弈信息. 而非完备信息博弈双方信息不完全透明,其中本文的研究背景就是假设零和博弈双方由于非完备信息博弈信息的不透明性导致其存在博弈诱导、欺诈等行为,即对抗性神经网络中的生成器与判别器之间的博弈对抗.

2 零和博弈对抗的 GAN 模型

自2012年深度学习的 AlexNet 模型问世以来,深度学习在很多领域的都取得了突破性进展^[2]. 2014年Goodfellow等人启发自博弈论中的二人零和博弈^[1,3-4],提出了生成对抗网络(Generative Adversarial Networks, GAN). 零和博弈是博弈论中的概念,是指一项对抗博弈中有输有赢,一方所赢正是另一方所输,而游戏的总成绩永远为零. 生成对抗网络包含一个生成器和一个判别器^[3,5-6]. 其中,生成器负责样本数据的分布,而判别器主要是判别输入是真实数据还是生成的样本. 这个模型是一个“极小极大博弈”问题^[7-9].

生成对抗网络(GAN)包含一个生成模型 G 和一个判别模型 D ,其结构如图1所示. 生成器和判别器两个网络彼此博弈^[10-11]. 生成器的目标是生成一个对象,并使其和真的一样,而判别器的目标就是找到生成出的结果和真实图像之间的差异^[3,12]. 在博弈对抗的过程中,由于判别器在训练后期无法辨别生成器产生的图像,进而生成器达到欺骗判别器的目的. 上述过程的数学抽象表示为生成式对抗网络会训练并更新判别器 D ,能将真实的分布 $P_{\text{data}}(x)$,和生成分布 $P_{\text{model}}(G(z))$ 区分开来. Z 表示属于某一分布的噪声数据,经过生成器后,变成 $X=G(z)$,得到生成分布 $P_{\text{model}}(G(z))$.

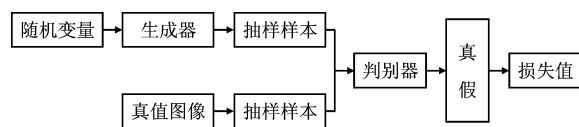


图1 对抗性神经网络结构图

算法模型最重要的评价指标之一是代价函数,对判别部分而言,所有的 GAN 变体都是使用相同的损失函数 $J(D)$ 定义,他们之间不同之处在于生成部分选择的代价函数是 $J(G)$.

先定义样本是从真实分布 $P_{\text{data}}(x)$ 来的,因此有:

$$E_{x \sim p_{\text{data}}} \log D(x) \quad (1)$$

式中, E 表示为期望,式子表示判别器 D 在 X 服从于 data 的概率密度时对数值的期望.

另一项是根据负类的对数损失函数构建的:

$$E_{z \sim p_{\text{model}}} \log(1 - D(G(z))) \quad (2)$$

综上所述,总的目标函数如下

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{data}}} [\log D(x)] + E_{z \sim p_z} [\log(1 - D(G(z)))] \quad (3)$$

式中, P_{data} 数据分布,通常以样本表示; P_z 是模型分布,其中 Z 通常建模为高斯模型.

3 对抗博弈的交叉熵代价函数

当“S”型函数的输出值接近1或者0时,网络权重和偏置项会出现梯度消失,识别精确率不高的情况,这时交叉熵函数在梯度消失的情况下仍然可以学习到数据,梯度存在且能保持较高的学习速率对网络参数进行更新,可以解决网络陷入局部最优问题.

相较于均值代价函数,在对抗性神经网络系统改进中将代价函数修改为交叉熵代价函数. 交叉熵代价函数采用形如式(4)所示的代价函数形式推出.

$$C = (y - a)^2 / 2 \quad (4)$$

式中, y 是期望输出即实际值, a 是神经元的输出, C 是输出的概率值, $a = \sigma(z)$, $z = wx + b$.

使用链式法则来求权重和偏置的导数如式(5)、(6)所示.

$$\frac{\partial C}{\partial W} = (a - y) \sigma'(z) x \quad (5)$$

$$\frac{\partial C}{\partial b} = (a - y) \sigma'(z) \quad (6)$$

通过消除公式(5)和(6)中的导数项 $\sigma'(z)$,则可以在激活函数接近或持续为0或者1时,梯度存在且能保持较高的学习速率对网络参数进行更新.

首先将式(5)和式(6)中的导数项去掉,可得式(7)、(8)所示的方程.

$$\frac{\partial C}{\partial w} = (a - y)x \quad (7)$$

$$\frac{\partial C}{\partial b} = (a - y) \quad (8)$$

联立式(4)、式(6)得:

$$\frac{\partial C}{\partial b} = (a - y)\sigma'(z) = \frac{\partial C}{\partial a}\sigma'(z) \quad (9)$$

由 $\sigma'(z) = \sigma(z)(1 - \sigma(z)) = a(1 - a)$, 可得:

$$\frac{\partial C}{\partial b} = \frac{\partial C}{\partial a}a(1 - a) \quad (10)$$

联立式(8)、式(10)得:

$$\frac{\partial C}{\partial a} = \frac{a - y}{a(1 - a)} \quad (11)$$

对式(11)左右两端以 a 为变量进行积分,可得单个神经元的交叉熵代价函数为:

$$C = -[y \ln a + (1 - y) \ln(1 - a)] + C_v \quad (12)$$

式中: C_v 是常数.

若是 n 个神经元进行训练,则所有的训练样本需进行平均,此时代价函数为:

$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] + C_v \quad (13)$$

综上同理可得对抗性神经网络交叉熵损失函数为:

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{data}}} [D(x) \log D(x)] + E_{z \sim p_z} [(1 - D(G(z))) \log(1 - D(G(z)))] \quad (14)$$

4 实验设计与论证

本文对抗样本取自于 MNIST 数据集,在损失函数改进基础上的对抗性神经网络进行二人零和博弈仿真.将本文提出的交叉熵损失函数算法在 Intel (R) Core (TM) i7-10750H, CPU 主频为 2.60 GHz,内存为 16GB 的硬件实验平台和软件 jupyter notebook 仿真平台上采用 Python 语言编程实现,并对算法的有效性进行分析研究.

4.1 仿真结果定量分析

首先应用均值损失函数对 MNIST 数据集的标准数字图像进行 28 000 次的训练迭代研究.图 2 是对抗性神经网络采用均值损失函数时,生成器与判别器的损失值对比图,即为 G_LOSS 和 D_LOSS.从仿真结果可以看出,生成器的损失值最高值为 4.7,之后数值逐步回归在 3.8 左右平稳变化;随着生成器的水平不断提高,判别器的损失值在降低,数值在 0.3 左右变化. G_

LOSS 和 D_LOSS 是逐渐收敛的,但是生成器与判别器之间的收敛区域过大,收敛的效果不佳.

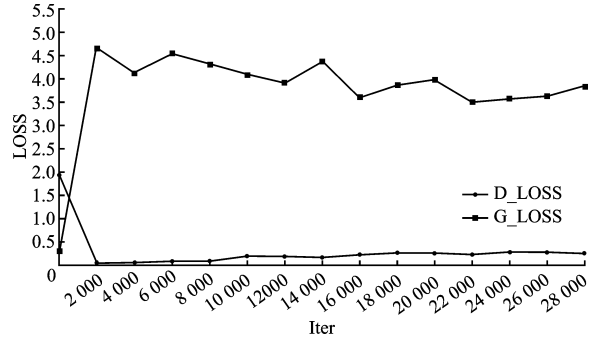


图 2 采用均值代价函数的生成器与判别器的损失值对比

其次应用交叉熵损失函数对 MNIST 数据集的标准数字图像进行 28 000 次的仿真迭代研究.图 3 是对抗性神经网络采用交叉熵损失函数时生成器与判别器损失值对比图像.即为 G_LOSS 和 D_LOSS.从仿真结果可以看出,生成器的损失值最高值为 8.5,之后数值逐步回归在 2.5 左右平稳变化,2 000~8 000 个周期期间损失值迅速降低,之后控制器将其值稳定控制在 2.5 左右;可以看出采用交叉熵损失函数的对抗性神经网络要比均值损失函数系统更能说明具有鲁棒性.随着生成器的水平不断提高,判别器的损失值在降低,数值在 0.5 左右变化. G_LOSS 和 D_LOSS 是逐渐收敛的,生成器与判别器之间的收敛区域相较于均值损失函数要小得多,说明采用交叉熵代价函数的对抗性神经网络系统中的生成器更具欺骗性,更有利于长期获得效益.

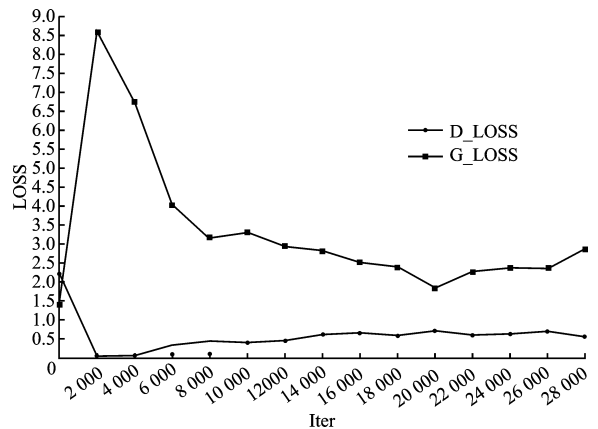


图 3 采用交叉熵代价函数的生成器与判别器的损失值对比

根据图 2、图 3 的迭代次数与损失值之间的关系,分别对应提取出开始训练阶段,损失值最低值节点,以及训练结束节点的均值代价函数以及交叉熵代价函数生成器对比数字图像.本文对比图像是在

jupyter notebook 的平台上生成,生成图像规格为 8×8 的图像,如图 4 所示。

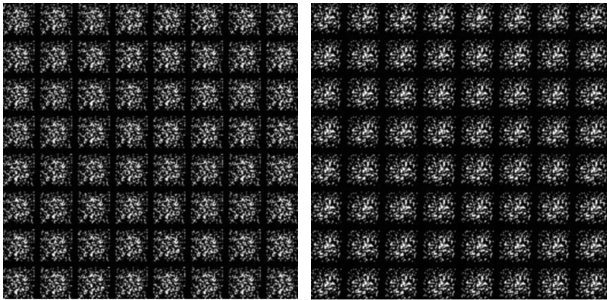


图 4 开始训练阶段的均值和交叉熵代价函数生成器对比图像

根据图 5、图 6 可以看出,在训练初始阶段,改进型代价函数并未显示出优势,但是随着训练的加深,改进型代价函数训练效果持续提升。当训练 Loss 值达到最低时,改进型代价函数产生的图像明显比均值代价函数清晰。在训练接近结束阶段,采用 OpenCV 和 Python 提取相应目标区域图像的像素值,均值和交叉熵代价函数生成器图像与原数据集图像相似度达到 70% 以上,即生成器生成图像像素点与原数据集图像像素点比值。图 5 中均值代价函数生成器图像像素点比值为 73.4%,交叉熵代价函数像素点比值 78.6%;图 6 中均值代价函数生成器图像像素点比值为 77.2%,交叉熵代价函数像素点比值 84.8%。

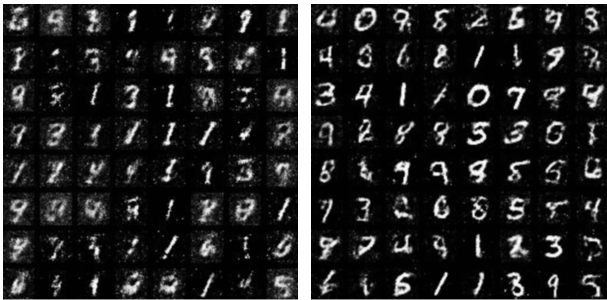


图 5 训练 Loss 最低值的均值和交叉熵代价函数生成器对比图像



图 6 训练结束点的均值和交叉熵代价函数生成器对比图像

4.2 仿真结果定性分析

为了避免上述实验数据及对抗性神经网络生成器的对比图像定量分析存在偶然性因素,所以加入数据的定性分析。表 1 是均值损失函数和交叉熵损失函数在训练迭代 28 000 次后的损失值。

表 1 损失函数的 G_Loss 和 D_Loss 对比

step	均值损失函数		交叉熵损失函数	
	G_Loss	D_Loss	G_Loss	D_Loss
1	0.295	1.921	1.400	2.226
2 000	4.653	0.032	8.571	0.008
4 000	4.123	0.041	6.721	0.080
6 000	4.547	0.077	4.017	0.330
8 000	4.314	0.135	3.144	0.422
10 000	4.087	0.188	3.300	0.396
12 000	3.902	0.183	2.932	0.437
14 000	4.382	0.158	2.787	0.615
16 000	3.598	0.221	2.516	0.648
18 000	3.863	0.263	2.372	0.569
20 000	3.974	0.255	1.834	0.702
22 000	3.496	0.219	2.253	0.589
24 000	3.569	0.273	2.363	0.628
26 000	3.624	0.270	2.356	0.686
28 000	3.847	0.248	2.849	0.550

由表格数据初步得知损失函数的 G_Loss 和 D_Loss 数据之间有着一定的相关性。这里采用 SPSS Statistics Version 19 做损失函数的 G_Loss 和 D_Loss 数据之间相关性分析。表 2 是对均值损失函数的 G_Loss 和 D_Loss 之间 Pearson 线性相关性分析;表 3 是均值损失函数的 G_Loss 和 D_Loss 之间 Kendall、Spearman 非线性相关性分析。

表 2 均值损失函数的 G_Loss 和 D_Loss 之间线性相关性分析

双变量线性相关性分析		G_Loss	D_Loss
G_Loss	Pearson 相关性	1	-0.972**
	显著性(双侧)		0.000
	N	15	15
D_Loss	Pearson 相关性	-0.972**	1
	显著性(双侧)	0.000	
	N	15	15

注:** 表示在 0.05 水平(双侧)上显著相关。

表 3 均值损失函数的 G_Loss 和 D_Loss 之间
非线性相关性分析

双变量非线性相关性分析		G_Loss	D_Loss
Kendall	相关系数	1.000	-0.611**
	G_Loss Sig. (双侧)		0.000
	N	29	29
	相关系数	-0.611**	1.000
	D_Loss Sig. (双侧)	0.000	
	N	29	29
Spearman	相关系数	1.000	-0.773**
	G_Loss Sig. (双侧)		0.000
	N	29	29
	相关系数	-0.773**	1.000
	D_Loss Sig. (双侧)	0.000	
	N	29	29

注：** 表示在置信度(双侧)为 0.01 时,相关性是显著的.

由上述表格不难分析出,均值损失函数的 G_Loss 和 D_Loss 的线性与非线性相关性分析系数的绝对值均大于 0.6. 即 $|r_{\text{Pearson}}| = 0.972 > 0.6$; $|r_{\text{Kendall}}| = 0.657 > 0.6$; $|r_{\text{Spearman}}| = 0.832 > 0.6$. 说明均值损失函数的 G_Loss 和 D_Loss 之间数据有强相关性,间接证明了数据定量分析是可信的.

同理可以得到交叉熵损失函数的 G_Loss 和 D_Loss 的 Pearson 线性相关性以及 Kendall、Spearman 非线性相关性分析表格,如表 4、表 5 所示.

表 4 交叉熵损失函数的 G_Loss 和 D_Loss 之间
线性相关性分析

双变量线性相关性分析		G_Loss	D_Loss
G_Loss	Pearson 相关性	1	-0.629*
	显著性(双侧)		0.012
	N	15	15
D_Loss	Pearson 相关性	-0.629*	1
	显著性(双侧)	0.012	
	N	15	15

注：* 表示在 0.05 水平(双侧)上显著相关.

同理由上述表格不难分析出,交叉熵损失函数的 G_Loss 和 D_Loss 的线性与非线性相关性分析系数的绝对值均大于 0.6. 即 $|r_{\text{Pearson}}| = 0.629 > 0.6$; $|r_{\text{Kendall}}| = 0.867 > 0.6$; $|r_{\text{Spearman}}| = 0.939 > 0.6$. 说明交叉熵损失函数的 G_Loss 和 D_Loss 之间数据有强相关性,间接证明了 4.1 节数据定量分析是可信的.

表 5 交叉熵损失函数的 G_Loss 和 D_Loss 之间
非线性相关性分析

双变量非线性相关性分析		G_Loss	D_Loss
Kendall	相关系数	1.000	-0.867**
	G_Loss Sig. (双侧)		0.000
	N	15	15
	相关系数	-0.867**	1.000
	D_Loss Sig. (双侧)	0.000	
	N	15	15
Spearman	相关系数	1.000	-0.939**
	G_Loss Sig. (双侧)		0.000
	N	15	15
	相关系数	-0.939**	1.000
	D_Loss Sig. (双侧)	0.000	
	N	15	15

注：** 表示在置信度(双侧)为 0.01 时,相关性是显著的.

由上述表格可以得出,均值损失函数偏向于线性相关性,而交叉熵损失函数则体现出非线性相关性,由 4.1 节的函数图像观察与表格分析相互验证这一结论. 至此,相关性分析指出 G_Loss 和 D_Loss 的变化是会互相影响指标因素. 因此在训练对抗样本时要充分体现生成器与判别器之间这种相辅相成的关系. 最终对抗性神经网络的改进型代价函数作用于生成器损失值降低,使得零和博弈优势充分体现.

5 结束语

本文在零和博弈模型对抗性神经网络中利用交叉熵代价函数算法替代了传统的均值损失函数算法,增强了生成器生成样本的效果,降低了判别器的能力. 充分拉近了生成器与判别器损失值之间的阈值区. 这将对博弈系统中生成器一方起到积极作用. 就本文而言,以极小的代价使得生成器模拟生成数据集中的图像来的对抗判别器的识别行为.

参考文献:

[1] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680.

[2] 张营营. 生成对抗网络模型综述[J]. 电子设计工程, 2018, 26(5): 34-37. DOI: 10.3969/j.issn.1674-6236.2018.05.007.

- ZHANG YY. A survey on generative adversarial networks[J]. Electronic Design Engineering, 2018, 26(5): 34-37. DOI: 10.3969/j.issn.1674-6236.2018.05.007.
- [3] KIMC I, KIMM, JUNGS, et al. Simplified Fréchet distance for generative adversarial nets[J]. Sensors, 2020, 20(6): 1548. DOI: 10.3390/s20061548.
- [4] LEE Y, CHO S. Design of semantic-based colorization of graphical user interface through conditional generative adversarial nets[J]. International Journal of Human-Computer Interaction, 2020, 36(8): 699-708. DOI: 10.1080/10447318.2019.1680921.
- [5] LI M Y, LIU Z Q, SHI X H, et al. ATCS: auto-tuning configurations of big data frameworks based on generative adversarial nets[J]. IEEE Access, 2020, 8: 50485-50496. DOI: 10.1109/ACCESS.2020.2979812.
- [6] CHEN C Q, XIE W P, WEN Y, et al. Multiple-source domain adaptation with generative adversarial nets[J]. Knowledge-Based Systems, 2020(199): 105962. DOI: 10.1016/j.knosys.2020.105962.
- [7] REN Y, LIS C, NIEM H, et al. Single image de-raining via improved generative adversarial nets[J]. Sensors (Basel), 2020, 20(6): 1591. DOI: 10.3390/s20061591.
- [8] 曹志义, 牛少彰, 张继威. 基于半监督学习生成对抗网络的人脸还原算法研究[J]. 电子与信息学报, 2018, 40(2): 323-330. DOI: 10.11999/JEIT170357.
- CAO Z Y, NIUS Z, ZHANG J W. Research on face reduction algorithm based on generative adversarial nets with semi-supervised learning[J]. Journal of Electronics and Information Technology, 2018, 40(2): 323-330. DOI: 10.11999/JEIT170357.
- [9] WANG W, LIU W D, LI J, et al. A rub fault recognition method based on generative adversarial nets[J]. Journal of Mechanical Science and Technology, 2020, 34(8): 1389-1397.
- [10] 林嘉豪, 章宗长, 姜冲, 等. 基于生成对抗网络的模仿学习综述[J]. 计算机学报, 2020, 43(2): 326-351. DOI: 10.11897/SP.J.1016.2020.00326.
- LIN J H, ZHANG Z C, JIANG C, et al. A survey of imitation learning based on generative adversarial nets[J]. Chinese Journal of Computers, 2020, 43(2): 326-351. DOI: 10.11897/SP.J.1016.2020.00326.
- [11] 许哲豪, 陈玮. 基于生成对抗网络的图片风格迁移[J]. 软件导刊, 2018, 17(6): 207-209. DOI: 10.11907/rjdk.172830.
- XU Z H, CHEN W. Style Transfer with generative adversarial nets[J]. Software Guide, 2018, 17(6): 207-209. DOI: 10.11907/rjdk.172830.
- [12] REN Y, LI S C, NIE M H, et al. Single image de-raining via improved generative adversarial nets[J]. Sensors (Basel), 2020, 20(6): 1591. DOI: 10.3390/s20061591.

作者简介:

卢向日 男,(1992-),博士研究生.研究方向为计算机控制、人工智能博弈.

汪湛清(通讯作者) 女,(1968-),高级工程师.研究方向为嵌入式系统与人工智能.

E-mail: bitwangzhanqing@163.com.

马宏宾 男,(1978-),教授.研究方向为人工智能与智能博弈.