

Text Generation in Civil Aviation Radiotelephony Communication Using Generative Adversarial Network

Yi Qiu*, Zihao Zhao, Dan Li, Fangyuan Cheng

Electric Information and Automation Department, Civil Aviation University of China, Tianjin
Email: *13072035081@163.com

Received: Dec. 7th, 2018; accepted: Dec. 20th, 2018; published: Dec. 27th, 2018

Abstract

Civil aviation radiotelephony communication is the main way for pilots and controllers to conduct voice communication. It is the information carrier that the controller commands to send and respond to the pilot's instructions. Pilots' understanding the order correctly by controllers is significant for the safety of civil aviation transportation. Using the deep learning method to perform automatic semantic verification in civil aviation radiotelephony communication can timely discover the contents of the pilot's read-back errors. Considering that training an effect automatic semantic verification model needs a lot of text data, this paper proposed a method of text generation in civil aviation radiotelephony communication using Generative Adversarial Network. First, the existing texts in civil aviation radiotelephony communication are classified and converted into one-hot vectors. Second, the new texts in civil aviation radiotelephony communication are generated based on the seq GAN model. And we compare the training effect of network based on different parameters. Finally, the generated texts are removed, which do not meet the grammar rules in civil aviation radiotelephony communication. This paper compared the effects of network training under different parameter selections. Experiment shows that GAN can be used to generate texts that comply with the civil aviation radiotelephony communication's rules.

Keywords

Land and Air Call Text, Generative Adversarial Network, Text Generation

基于GAN的民航陆空通话文本生成方法

邱意*, 赵子豪, 李丹, 程方圆

中国民航大学电子信息与自动化学院, 天津
Email: *13072035081@163.com

*通讯作者。

收稿日期：2018年12月7日；录用日期：2018年12月20日；发布日期：2018年12月27日

摘要

民航陆空通话是飞行员与管制员进行话音通信的主要方式，是管制员指令发送与飞行员指令回复的信息载体，飞行员正确理解管制员所发出的指令对于民航运输安全有着重要意义。采用深度学习方法对陆空通话进行实时语义校验可以及时发现飞行员错误的复诵内容。考虑到训练一个有效的差错校验网络模型需要大量的文本数据，本文提出一种基于生成对抗网络GAN的陆空通话文本生成方法。首先对现有真实的陆空通话文本进行筛选和分类，并将其转换成one-hot词向量。其次，使用one-hot训练GAN模型，并使用其生成新的陆空通话文本。最后，根据陆空通话语法规则筛选出符合要求的生成文本。本文比较了不同参数选择下网络训练的效果，通过实验验证，使用GAN模型能生成符合民航陆空通话语言规则的文本。

关键词

陆空通话文本，生成式对抗网络，文本生成

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

民航安全在中国民航事业的发展中具有重要的作用与意义。在民航运输健康运行中，为了保证航空器安全、高效的运行，管制员和飞行员必须清晰准确的理解双方的指令内容，进而保证导航指令准确无误地进行传递。民航陆空通话语音指令的识别，语义一致性校验的技术在保证航空器飞行安全中具有重要作用[1] [2] [3]。因此，搭建可靠的民航陆空通话语义感知和智能校验系统对于自动检测陆空通话复诵中的误差具有辅助作用。智能校验系统不仅可以自动检测管制员与飞行员语义的一致性，还可以使相关人员的用语更规范，从而提高民航运输的安全系数。民航陆空通话智能校验系统[4]实现的主要步骤为：1) 构建符合陆空通话规范的文本语料库；2) 搭建适用于语义一致性检验深度学习的框架；3) 训练语义一致性智能校验模型；4) 验证模型的可靠性。大规模的陆空通话语料库是后期训练模型的基础，但目前并没有公开、公认的陆空通话语料库供研究使用，都是利用人工方式构建而成。在已有陆空通话文本语料库进行模型训练过程中存在陆空通话文本语料库文本较少、文本多样性不足的问题，这会导致深度学习模型容易过拟合、泛化能力不够好。因此如何获得大规模、高质量的符合陆空通话规范的标准语料库是亟待解决的问题。

现有的扩大数据方式一般是人工抽取和标注，由于人工抽取和标注的方式工作量大且易出错，使用机器代替人工进行语料库的扩充具有重大的研究意义。近些年，生成对抗网络(Generative Adversarial Networks, GANs)在数据生成上表现出很好的性能，提供了一种不需要大量标注训练数据就能学习深度表征的方式，从而可以产生与真实数据相同分布的海量数据。

生成对抗网络目标旨在构成一个生成器，它将来自给定先验分布的样本功能性地映射到合成数据[5]，其核心思想是构建两个神经网络：判别器(Discriminator, D)和生成器(Generator, G)，生成器不断

的生成假样本来欺骗判别器，判别器判断样本是真样本和假样本；判别器和生成器交替训练和优化，使生成样本更加真实，达到生成海量数据的目的。目前 GAN 已经在图像生成[6] [7]和文本生成[8] [9] [10]上有了很多的应用。文章[11]中提出 GAN 虽然在计算机视觉领域取得巨大的成功，但对于离散型数据的文本序列却捉襟见肘。文章[12]以长短时记忆网络(Long Short Time Memory Network, LSTM)作为 GAN 的生成器，以卷积神经网络(Convolutional Neural Networks, CNN)作为 GAN 的判别器，并使用光滑近似的思想逼近生成器 LSTM 的输出，从而解决离散导致的梯度不可导问题。文章[13]对 GAN 的损失函数做出修改，对真实句子和生成句子的隐藏特征表示进行匹配操作，这种方法可以得到更好的收敛效果。

陆空通话语言自动校验任务面临着数据量小，模型训练不充分的问题。针对这个问题，本文提出陆空通话文本生成 GAN 模型来扩充样本并提高多样性，减少人工标注的繁琐。首先，对真实的民航陆空通话文本进行序列长度整理，并对文本的词个数进行统计，提出每个词的 one-hot 向量，然后，使用递归神经网络(Recurrent Neural Network, RNN)搭建文本生成器，使用 CNN 搭建文本判别器，并初始化它们；在规定的训练次数中，使用真实数据训练生成器与判别器；最后，在对抗训练中，使用判别器对生成器进行动态调整训练。

2. 民航陆空通话文本生成模型

本文提出的基于 GAN 的民航陆空通话文本生成模型框架如图 1 所示，该模型由两部分组成：文本生成器，文本判别器。

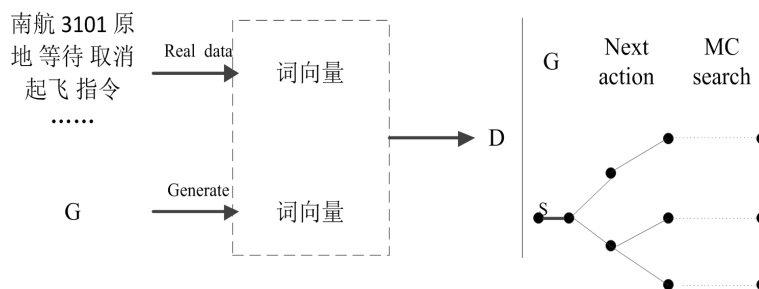


Figure 1. The GAN model of civil aviation radiotelephony communication
图 1. 民航陆空通话文本生成 GAN 模型

2.1. 文本生成器

在民航陆空通话中，为保证飞行员与管制员的通话语义清晰，通话的规则有以下特点：1) 单词发音特殊，如，1 读作幺，A 读作 ALPHA；2) 通话用词固定，如进近：approach，终止：terminate；3) 句式固定，飞行员与管制员之间严格按照“对方 - 己方 - 内容”的固定语法通话，如飞行员(Beijing TWR, CCA123.Request start up.)，管制员(CCA123, Beijing TWR, Start up approved.)。

在民航陆空通话文本生成任务中，生成器的目标是实现文本序列的生成。本文使用 RNN [14]搭建生成模型。在生成模型中，RNN 网络以输入序列 $X = (x_1, x_2, \dots, x_T)$ 的 one-hot 向量为输入，通过 RNN 网络中隐层的语义提取，得到隐层中包含语义信息的状态序列 h_1, \dots, h_T ，其计算公式如下(1)所示。

$$h_t = f(h_{t-1}, x_t) \tag{1}$$

然后，生成模型中 softmax 层以状态序列为输入，输出输入序列 X 的输出标签 $Z(z_1, z_2, \dots, z_m)$ 概率分布，计算公式如公式(2)表示。

$$p(z_t | x_1, \dots, x_t) = \text{softmax}(c + Vh_t) \quad (2)$$

其中, C 表示偏置量, V 表示权值矩阵。

2.2. 文本判别器

在民航陆空通话文本生成中, 文本生成器训练需要文本判别器参与, 从而生成特定的文本。对于文本生成器生成的结果序列, 文本判别器的功能是判别输入的文本序列 X 是否为真实数据, 并通过调整目标函数对生成器进行动态更新。随着深度学习的快速发展, 在文本分类任务中, 应用深层的 CNN 网络与 RNN 网络模型能取得令人满意的效果。本文使用 CNN 搭建文本判别器, 并使用它判别生成器生成的结果序列是否为真实的民航陆空通话文本序列。

对于给定的输入序列 $X = (x_1, x_2, \dots, x_T)$, 本文首先对 X 进行编码, 如公式(3)所示。

$$\chi_{1:T} = x_1 \oplus x_2 \oplus \dots \oplus x_T \quad (3)$$

其中 x_t 表示一个 k 维的词向量, 且 $x_t \in \mathbb{R}^k$, \oplus 表示词向量间的连接处理。 $\chi_{1:T}$ 表示词向量的连接矩阵, 且 $\chi_{1:T} \in \mathbb{R}^{T \times k}$ 。对于词向量矩阵, 然后使用一个长度为 l 的卷积核 w_j 进行卷积, 且 $w \in \mathbb{R}^{l \times k}$, 得到一个新的特征图 c_i , 其中计算入公式(4)所示。

$$c_{ji} = \varphi(w_j \otimes \chi_{ii+l-1} + b) \quad (4)$$

其中 \otimes 表示卷积核与词向量矩阵进行卷积求和运算, φ 表示一个非线性函数。对于卷积神经网络, 不同大小的卷积核将会提取不同的民航陆空通话文本语义特征。因此, 本文参考[4], 选取最大池化, 如公式(5)所示。

$$c_{ji} = \max\{c_{j1}, \dots, c_{jT-l+1}\} \quad (5)$$

最后在文本生成器中, 提取的特征通过一个激活函数为 sigmoid 函数的全连接层, 得到输入文本序列 X 是真实民航陆空通话文本序列的概率值, 其中 sigmoid 函数形式如公式(6)所示。本文在文本生成器的 CNN 网络训练中, 使用交叉熵损失函数(7)作为目标函数。

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

$$\min\{\text{loss}(\log D(Z)) - \text{loss}(\log(1 - D(Z)))\} \quad (7)$$

2.3. 模型梯度训练方法

本文参考[15]中模型训练方法, 首先使用文本生成器生成一个文本序列, 并求输出文本序列的累积期望值 $J(\theta)$, 计算公式如公式(8)所示。

$$J(\theta) = \sum_{Z_{1:T}} G_\theta(Z_{1:T} | X) \cdot R_D^{G_\theta}(Z_{1:T-1}, X, z_T, Z^*) \quad (8)$$

其中, θ 表示生成器 G 的参数, $Z_{1:T}$ 表示 G 生成的文本序列, X 表示真实的民航陆空通话文本序列, Z^* 表示期望的生成民航陆空通话文本序列。 $R_D^{G_\theta}$ 是一个函数, 对于给定的 G_θ , 它可以计算文本序列 X 在当前状态 $s(Z_{1:T-1}, X)$, 转移到下一个状态的概率 $s(Z_{1:T}, X)$ 。文本生成任务的期望是能生成一个被文本判别器判别为真实数据的文本序列, 在本文中, 使用文本判别器 $D(Z_{1:T}^n)$ 的值作为第 n 次搜索的结果序列 $Z_{1:T}^n$ 是真实数据的概率值。但是在实际中, 如果 $Z_{1:T}^n$ 不是一个完整的民航陆空通话文本序列, 本文将使用蒙特卡洛树搜索(MCTS)代替文本判别器 D 的输出值, 并使用其输出值训练文本生成器 G , 在 N 次的 MCTS 搜索中有:

$$\{Z_{1:T}^1, \dots, Z_{1:T}^N\} = \text{Mc}^{G_\theta}((Z_{1:t}, X), N) \quad (9)$$

其中 $Z_{1:T}^N$ 表示在时间 T 内, 由 MCTS 生成的第 N 个序列。 $(Z_{1:t}, X) = (z_1, z_2, \dots, z_t, X)$ 表示当前搜索状态, $Z_{t+1:T}^N$ 表示由 G_θ 生成的词。综上所述有:

$$R_D^{G_\theta}(Z_{1:T-1}, X, z_T, Z^*) = \begin{cases} \frac{1}{N} \sum_{n=1}^N D(Z_{1:T}^n), Z_{1:T}^n \in \text{Mc}^{G_\theta}((Z_{1:t}, X), N) & \text{for } t < T \\ D(Z_{1:t}) & \text{for } t = T \end{cases} \quad (10)$$

在文本生成器 G_θ 的训练中, 使用文本判别器 D 的输出值可以参与 G_θ 的迭代更新中, 并且可以使用生成的真实文本序列重新训练文本判别器 D , 如公式(10)所示。

$$\min \left\{ \text{loss}_{Z^* \in P_{data}}(\log D(Z)) - \text{loss}_{Z^* \in G_\theta}(\log(1 - D(Z))) \right\} \quad (11)$$

在更新文本判别器 D 后, 将继续对文本生成器 G_θ 的权值 θ 进行更新, 其计算公式如(11)所示:

$$\begin{aligned} \nabla J(\theta) &= \frac{1}{T} \sum_{t=1}^T \sum_{Z_t \in Z} R_D^{G_\theta}(Z_{1:T-1}, X, z_T, Z^*) \cdot \nabla_\theta (G_\theta(z_t | Z_{1:t-1}, X)) \\ &= \frac{1}{T} \sum_{t=1}^T \text{loss}_{z_t \in G_\theta} \left[R_D^{G_\theta}(Z_{1:T-1}, X, z_T, Z^*) \cdot \nabla_\theta \log p(z_t | Z_{1:t-1}, X) \right] \end{aligned} \quad (12)$$

2.4. 模型训练

生成式对抗网络的训练是一个生成器与判别器博弈的过程, 需要生成器与判别器同时参与训练并更新参数, 其训练过程具有不稳定的特点。在本文中, 首先使用最大似然估计为目标函数对文本生成器进行预训练, 并每隔一定的训练次数, 使用生成器生成一部分数据, 且以最大交叉熵函数为目标函数, 使用这部分数据预训练判别器。在对抗训练中, 1) 首先本文使用训练完成的文本生成器生成民航陆空通话文本序列, 并使用预训练完成的文本判别器 D 与 MCTS 搜索方法, 对生成的文本计算 $R_D^{G_\theta}$ 值; 2) 然后更新文本生成器参数; 3) 最后, 在规定训练次数中使用更新的文本生成器生成文本, 并使用生成文本训练文本判别器。

3. 实验结果与分析

3.1. 实验数据

本文处理的数据集为中文陆空通话文本, 文本包含 24,780 条指令文本, 从中挑选 1000 条作为测试集, 10,000 条作为训练集。由于陆空通话的结构特殊, 一条指令文本包含航班呼号, 管制区号, 指令动作与动作状语, 如 CCA3818 北京塔台下降到 1500 米, 在其中包含中文, 英文及数字, 因此本文使用字符级别的 one-hot 向量作为词向量。

3.2. 参数设置

本文在预训练中, 对 RNN 网络搭建的文本的生成器 G 和 CNN 搭建的文本判别器 D 进行 20 次预训练。然后, 在 40 次对抗训练中, 本文参考论文[15][16], 使用两个方法对文本生成结果进行分析: 1) 使用 LSTM 网络搭建文本生成器 G_{lstm} , 将其参数初始化在[0,1], 并使用其对生成的民航陆空通话文本进行 Nll_{test} (Negative log-likelihood)值计算, 其值可以体现文本生成器生成真实文本的性能:

$$\text{Nll}_{\text{test}} = -\text{loss}_{Z_{1:T} \sim G_{\text{LSTM}}} \left[\sum_{t=1}^T \log G_\theta(z_t | Z_{1:t-1}) \right] \quad (13)$$

2) 为计算两个文本的相似性,对每次生成文本的词计算 EmbSim 值,其中 W_i' 表示生成的民航陆空通话文本中序列词的相似性,这里是计算两个单词词向量的余弦值,同样, W_i 表示真实民航陆空通话文本中词与词间的相似余弦值, N 表示文本词的个数,本文 N 取 10,000。整个实验中,对抗训练为 100, MCTS 搜索长度为 16。

$$\text{EmbSim} = \log \left(\sum_{i=1}^N \cos(W_i', W_i) / N \right) \quad (14)$$

3.3. 实验结果

在本文的实验中,本文使用不同的训练策略对文本生成器 G 与文本判别器 D 的训练进行训练。由于民航陆空通话的语言规则固定,为提高文本判别器 D 的性能,在预训练中,对 D 进行了不同次数的训练,并对抗训练中,使用它指导文本生成器 G 的训练。其实验结果如图 2 所示。

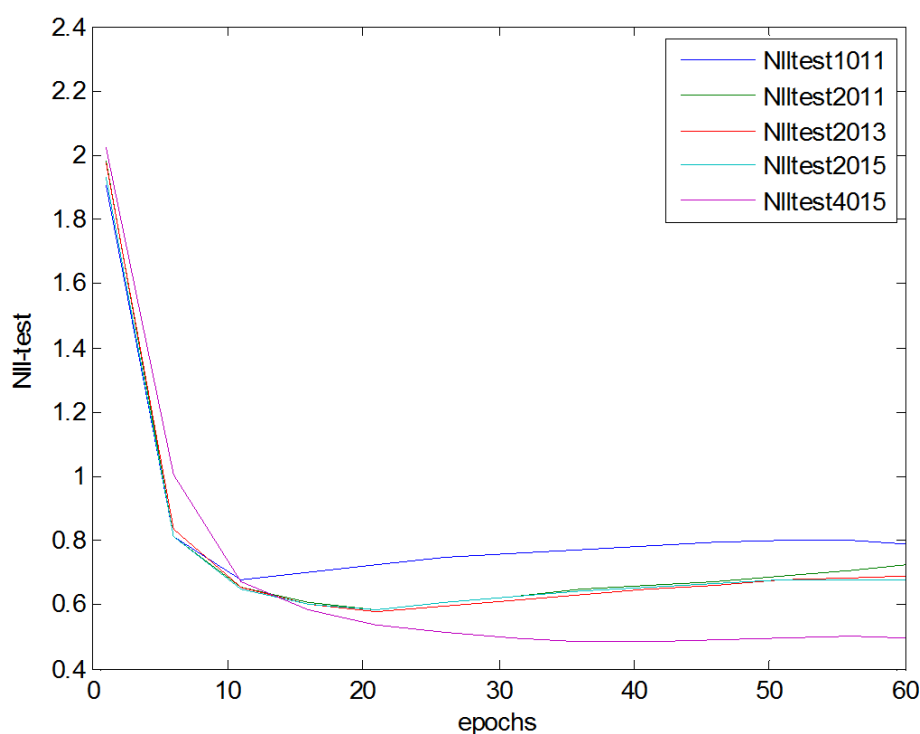


Figure 2. The Nll-test value results of different training policy

图 2. 不同训练策略的 Nll-test 实验结果

其中 Nlltest4015 表示预训练次数为 40 次,对抗训练中,文本生成器 G 训练 1 次,文本判别器 D 训练 5 次。从图 2 可知,不同的预训练次数影响 GAN 的民航陆空通话文本生成效果,随着预训练次数的增加,Nll-test 值越低,生成文本质量越好。因此,本文在后期实验中,在进行对抗训前,对生成器与判别器进行 40 次的预训练。在对抗训练中,生成器的更新需要判别器参与,从图 3 可知,当生成器与判别器进行同步训练时,GAN 的训练稳定。

在自然语言处理领域, BLEU 是一种广泛被使用的评价标准,可以对句子和文本中的词的相似性进行计算,得出一个 BLEU [17]值。对于本文生成的文本,将被使用在民航陆空通话语音识别的 3-gram 和 4-gram 语言模型训练中,因此,实验中选取 BLEU 的值为 2、3、4,实验结果如表 1 所示。本文训练的 GAN 模型在训练集与测试集上可以取得较好的 BLEU 值,生成的文本符合民航陆空通话规则。

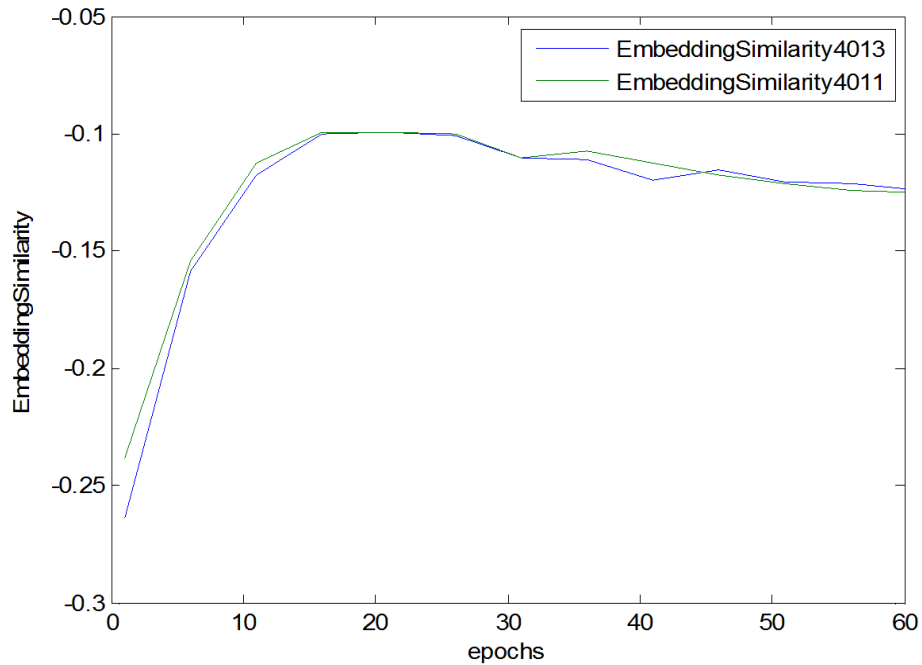


Figure 3. The EmbSim value results of different training policy
图 3. 不同训练策略的 EmbSim 实验结果

Table 1. The BLEU value of training dataset and test dataset
表 1. 训练集与测试集的 BLEU 结果

BLEU-*	Training data	Test data
2	0.924	0.814
3	0.811	0.513
4	0.542	0.486

4. 结论

在自然语言处理中，GAN 是实现无监督学习与样本生成的重要方法，使用 GAN 生成文本序列是自然语言理解的高级阶段。在为实现民航陆空通话的文本生成中，本文探究了 GAN 在民航陆空通话文本生成中的应用，从实验结果分析可得，GAN 生成的民航陆空通话文本在句子结构方面能与真实的民航陆空通话文本相同，但是存在个别生成序列不符合民航陆空通话语言规则，如(国航 4811 近地告警上升到修正海压向南)，对于真实的民航陆空通话文本，“修正海压”后应跟具体数值，如 1010，而非“向南”。但总体而言，使用 GAN 生成的民航陆空通话文本基本符合实际数据，有利于进行陆空通话语义校验的研究。

致 谢

感谢中央高校基本科研业务费资助项目(项目编号: ZYGX2018042)的资助。

基金项目

中央高校基本科研业务费资助项目，项目编号: (ZYGX2018042)。

参考文献

- [1] 王万乐, 赵琦. 空中交通管制员无线陆空通话[M]. 北京: 北京交通大学出版社, 2016.
- [2] 董昱廷. 陆空通话对飞行安全影响的分析[J]. 移动信息, 2016(6): 00162-00163.
- [3] Cui, Q. and Li, Y. (2015) The Change Trend and Influencing Factors of Civil Aviation Safety Efficiency: The Case of Chinese Airline Companies. *Safety Science*, **75**, 56-63.
- [4] Jia, G.M., Lu, Y.J., Lu, W.B., *et al.* (2017) Verification Method for Chinese Aviation Radiotelephony Readbacks Based on LSTM-RNN. *Electronics Letters*, **53**, 401-403. <https://doi.org/10.1049/el.2016.2877>
- [5] Goodfellow, I., Pouget-Abadie, J., Mirza, M., *et al.* (2014) Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, 2672-2680.
- [6] Isola, P., Zhu, J.-Y., Zhou, T.H. and Efros, A.A. (2017) Image-to-Image Translation with Conditional Adversarial Networks. arXiv:1611.07004 [cs.CV]
- [7] Taigman, Y., Polyak, A. and Wolf, L. (2017) Unsupervised Cross-Domain Image Generation. arXiv:1611.02200 [cs.CV]
- [8] Che, T., Li, Y., Zhang, R., *et al.* (2017) Maximum-Likelihood Augmented Discrete Generative Adversarial Networks. arXiv:1702.07983 [cs.AI]
- [9] Lin, K., Li, D., He, X., *et al.* (2017) Adversarial Ranking for Language Generation. arXiv:1705.11001 [cs.CL]
- [10] Rajeswar, S., Subramanian, S., Dutil, F., *et al.* (2017) Adversarial Generation of Natural Language. arXiv:1705.10929 [cs.CL]
- [11] Kusner, M.J. and Hernández-Lobato, J.M. (2016) GANS for Sequences of Discrete Elements with the Gumbel-softmax Distribution. arXiv:1611.04051 [stat.ML]
- [12] Zhang, Y., Gan, Z. and Carin, L. (2016) Generating Text via Adversarial Training. https://zhegan27.github.io/Papers/textGAN_nips2016_workshop.pdf
- [13] Zhang, Y., Gan, Z., Fan, K., *et al.* (2017) Adversarial Feature Matching for Text Generation. arXiv:1706.03850 [stat.ML]
- [14] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [15] Yu, L., Zhang, W., Wang, J., *et al.* (2016) GAN: Sequence Generative Adversarial Nets with Policy Gradient. arXiv:1609.05473 [cs.LG]
- [16] Zhu, Y., Lu, S., Zheng, L., *et al.* (2018) Texygen: A Benchmarking Platform for Text Generation Models. arXiv:1802.01886 [cs.CL]
- [17] Papineni, K., Roukos, S., Ward, T., *et al.* (2002) Bleu: A Method for Automatic Evaluation of Machine Translation. *ACL'02 Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia, 7-12 July 2002, 311-318.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2161-8801, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: csa@hanspub.org