

基于边缘计算中极端姿态和表情的人脸识别

况朝青^{1,2,3}, 贺超^{1,2,3}, 王均成^{1,2,3}, 邹建纹^{1,2,3}

(1.重庆邮电大学 通信与信息工程学院, 重庆 400065; 2.重庆高校市级光通信与网络重点实验室, 重庆 400065;
3.泛在感知与互联重庆市重点实验室, 重庆 400065)

摘要: 随着信息技术的发展, 人脸识别在支付、工作和安防系统中应用的越来越多。在边缘计算系统中, 为了处理的速度, 通常选择较小的神经网络进行人脸识别, 这样会导致识别率低。并且在实际应用中大多都是对于图片质量较高的人脸可以很好地识别, 但对于受光照影响较大、表情和姿态变化大的图片识别率不是很高。因此, 选择 SqueezeNet 轻量级网络, 该网络层数小, 可以很好地运用于边缘计算系统中。采用了预处理的方法来对图片进行预处理, 然后改进了 SqueezeNet 网络的损失函数以及加入了 ResNet 网络中的残差学习方法。最后通过对 LFW 和 IJB-A 数据集进行测试, 该研究方法明显提高了识别率。

关键词: 神经网络; 人脸识别; 预处理; SqueezeNet 网络; ResNet 网络

中图分类号: TN911.73; TP391.4

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.200968

中文引用格式: 况朝青, 贺超, 王均成, 等. 基于边缘计算中极端姿态和表情的人脸识别[J]. 电子技术应用, 2021, 47(6): 30-34.

英文引用格式: Kuang Chaoqing, He Chao, Wang Juncheng, et al. Face recognition with extreme posture and expression[J]. Application of Electronic Technique, 2021, 47(6): 30-34.

Face recognition with extreme posture and expression

Kuang Chaoqing^{1,2,3}, He Chao^{1,2,3}, Wang Juncheng^{1,2,3}, Zou Jianwen^{1,2,3}

(1.School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications,
Chongqing 400065, China;

2.Optical Communications and Networks Key Laboratory of Chongqing, Chongqing 400065, China;

3.Ubiquitous Sensing and Networking Key Laboratory of Chongqing, Chongqing 400065, China)

Absrtact: With the development of information technology, face recognition is used more and more in payment, work and security system. In the edge computing system, in order to deal with the speed, we usually choose a smaller neural network for face recognition, which may cause the recognition rate is not very high. And in practical applications, most of them can recognize the face with high image quality, but the recognition rate is not very high for the face which is greatly affected by the light and has great changes in expression and posture. Therefore, this paper chooses the SqueezeNet lightweight network, which has a small number of layers and can be well used in edge computing system. The method of preprocessing is used to preprocess the image, and then the loss function of SqueezeNet network and the residual learning method of ResNet network are improved. Finally, through the test of LFW and IJB-A data set, it is concluded that the research method in this paper can significantly improve the recognition rate.

Key words: neural network; face recognition; preprocessing; SqueezeNet network; ResNet network

0 引言

近年来, 人脸识别受到越来越多的关注, 主要是通过神经网络模型来进行人脸识别。但人脸识别依然是一个非常重要但又极具挑战性的问题, 主要是现在大部分的人脸识别采用的图像都是静态和质量较高的图片, 所以识别效果很好。但在实际应用中, 人脸图像受到光照、表情和较大的姿态变化的影响, 可能导致识别率急剧下降。因此, 采用一种预处理的方式来处理图片, 提高图片的质量, 成为了当下研究的关键^[1]。并且在边缘计算系

统中, 采用大型网络来进行人脸识别是不现实的, 主要是受到处理器的速度和功耗的影响, 因此这方面的应用成为了研究的热点。

针对这些问题, 文献[2]–[3]主要是通过位置向量的特征来解决人的姿势问题; 文献[4]–[5]则试图合成一个正面的人脸; 文献[6]提出了采用主成分分析(PCA)的方法, 主要是可以将样本数据向一个新的低维空间投影, 经过坐标转换后, 使用原样本中最大的一个线性无关特征值对应的空间坐标表示, 可以一定程度地去除样本噪

声;文献[7]用滤波器方法来提高识别率。

本文采用了神经网络的方法来对输入图片的人脸进行预处理,将预处理后生成的图片输入到网络中进行测试。本文采用在 SqueezeNet 网络的基础上加入残差网络,并对传统的 softmax 进行了修改。SqueezeNet 网络是一个轻量级网络,网络的层数不多,非常适合于边缘计算的识别系统。但如果直接进行人脸识别测试,可能识别率不是很高。因此,本文对图片加了预处理,然后对网络进行了适当改进。并且对网络改进以后,通过计算和测试,并没有增加网络的计算量。最后通过实验表明,该方法提高了识别的准确率。

1 预处理

人脸识别系统中很重要的一步就是人脸图像的预处理过程。本文主要采用神经网络的人脸归一化方法。人脸归一化是在保持人脸身份的前提下,从一张人脸图像中合成一张正态人脸。本文主要是将边缘计算中提取的人脸图片经过 MTCNN^[8]神经网络来对人脸图片进行裁剪,减少图片上的无用信息,使得提取的人脸的特征值更多。然后将裁剪后的图片通过文献[1]提出的方法进行处理,生成一张新的归一化后的人脸图片,图片的分辨率得到提高。

2 网络架构

FaceNet^[9]是一个通用的人脸识别系统,采用的方法是通过卷积神经网络将人脸图像映射到欧几里得空间,检测映射不同点之间的距离进而判断是否是同一个人。因为每个人都可以看成一个类,同一个人的不同人脸之间的距离可被看作类内距离,不同人的人脸图像之间的距离可被看作是类间距离。FaceNet 系统的主要思想可看作尽量缩小类内距离,增大类间距离。本文主要采用 FaceNet 网络架构,使用 SqueezeNet^[10]网络。然后对 SqueezeNet 网络进行了适当的改进,加入了残差网络,将损失函数改为了现在识别效果更好的损失函数。FaceNet 的网络结构如图 1 所示。

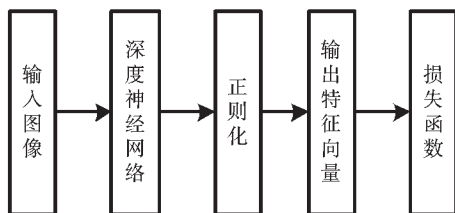


图 1 FaceNet 网络架构

2.1 网络模型

为了使模型适用于边缘计算系统中,并且能有一个好的识别率,本文对 SqueezeNet 网络架构进行了修改,并加入了预处理的过程,整个识别过程如图 2 所示。主要是将边缘计算中提取的原始人脸图像输入网络中先通过 MTCNN 网络进行裁剪人脸,再进行预处理,最后通过输入本文改进的神经网络结构,识别是否是同一个人。

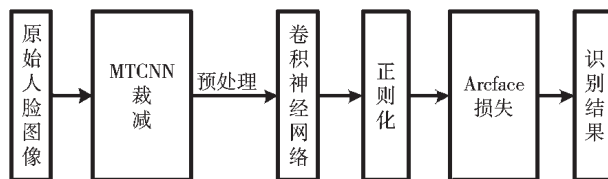


图 2 总的网络结构图

2.2 损失函数

本文采用的损失函数是 Arcface^[11-12]损失函数。Arcface 是对传统的 softmax^[13]的改进,传统的损失函数如式(1)所示,改进以后的损失函数如式(2)所示。Arcface 为了使类内对象紧致,类间对象分离,主要是对归一化后输入的 x (特征)和归一化后的 W (权重)通过叉乘得到一个 Arccos 的值,这个值对应的角度为 θ ;然后 m 加上 θ 后,将新的 $\cos(\theta+m)$ 与 s (特征缩放)相乘后取 \log ,再经过 softmax 输出。Arcface 的实现方式与 softmax 比较也没有变得更加复杂,但考虑了角度对分类的影响,通过验证能更好地提高识别率。

$$L_1 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{w_i^T x_i + b_i}}{\sum_{j=1}^n e^{w_j^T x_i + b_j}} \quad (1)$$

$$L_2 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{S(\cos(\theta_i + m))}}{e^{S(\cos(\theta_i + m))} + \sum_{j=1, j \neq y_i}^n e^{S \cos \theta_j}} \quad (2)$$

2.3 SqueezeNet 与 ResNet

SqueezeNet 的核心在于 fire 模块,fire 模块由两层构成,分别是 squeeze 层+expand 层。squeeze 层是一个 1×1 卷积核的卷积层;expand 层是 1×1 和 3×3 卷积核的卷积层,expand 层中,把 1×1 和 3×3 得到的特征图进行拼接(concat),具体操作如图 3 所示。Kernel 是卷积核大小,Num 是卷积核个数; H 、 W 、 M 分别是特征的长、宽、维数。

ResNet^[14]的架构如图 4 所示。通过这种残差网络结构,不仅可以在边缘计算中减少网络结构的计算量,还

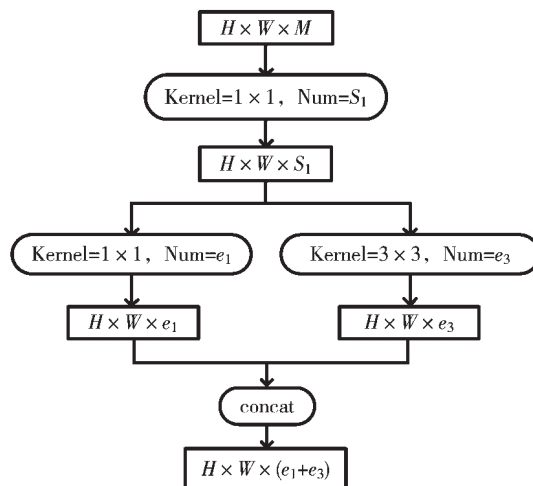


图 3 fire 模块

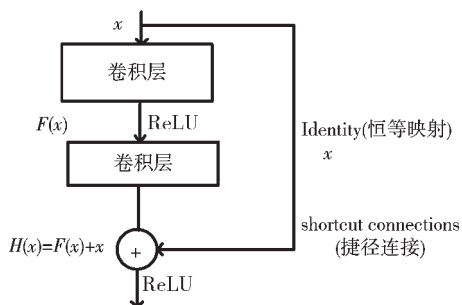


图4 残差网络模块

能提高识别率。残差网络用了高速网络^[15]的跨层链接思想,通过捷径连接(shortcut connections)的方式,直接把输入 x 传到输出作为初始结果,输出结果为 $H(x)=F(x)+x$,当 $F(x)=0$ 时,那么 $H(x)=x$,也就是上面所提到的恒等映射。

SqueezeNet 网络与加入残差模块后的网络架构如图5所示。本文主要测试了两种方式的残差模块,第一种是在隔一个的 fire 模块中加入残差网络,而第二种是在每一层中都加入残差模块。最后在识别中发现,第一种更好,并且不会增大网络的计算量,因此第一种改进的算法更适合用于边缘计算系统中。整个 SqueezeNet 就是使用 Fire 基本模块堆积而成的,网络结构如图5所示,其中左图是标准的 SqueezeNet,其开始是一个卷积层 conv1,后面是 Fire 模块的堆积,中间穿插着 stride=2 的最大池化(maxpool)层,其主要作用是下采样,并且采用延迟的策略,尽量使前面层有较大的特征图,最后经过池化(global avgpool)。中图 SqueezeNet+simple ResNet 和右图 SqueezeNet+complex ResNet 分别引入了不同的短路机制的 SqueezeNet,这是借鉴了 ResNet 的结构。

3 实验结果与分析

本实验所使用的计算机配置: Intel 的 i5 处理器, GTX1660Ti 的 GPU, Win10 操作系统。在 Anconda+TensorFlow 环境下运行。

3.1 预处理后的结果

首先对本文采用的预处理方法进行验证,图6显示了在极端姿态、表情和光照条件下该方法处理后的图片与原图的对比图。从上往下依次是姿态、表情、光照、遮挡、正常人脸的测试图;每行图中,两列图像为一组,一共展示了3组;其中每一组中第一列是原始图像,第二列是处理以后的图像。可以明显看出,将裁剪后的图片进行预处理后,人脸的分辨率得到了提高。

3.2 数据集和评价指标

本文主要采用 LFW 数据集和 IJB-A 数据集进行了测试。LFW 数据集的人脸图像大多数是正面的照片,没有干扰;而 IJB-A 数据集是在自然环境下拍摄的,因此受到姿态、光照和表情等的影响。

记输入的测试对为 (x_i, x_j) , 所有的同一身份测试对

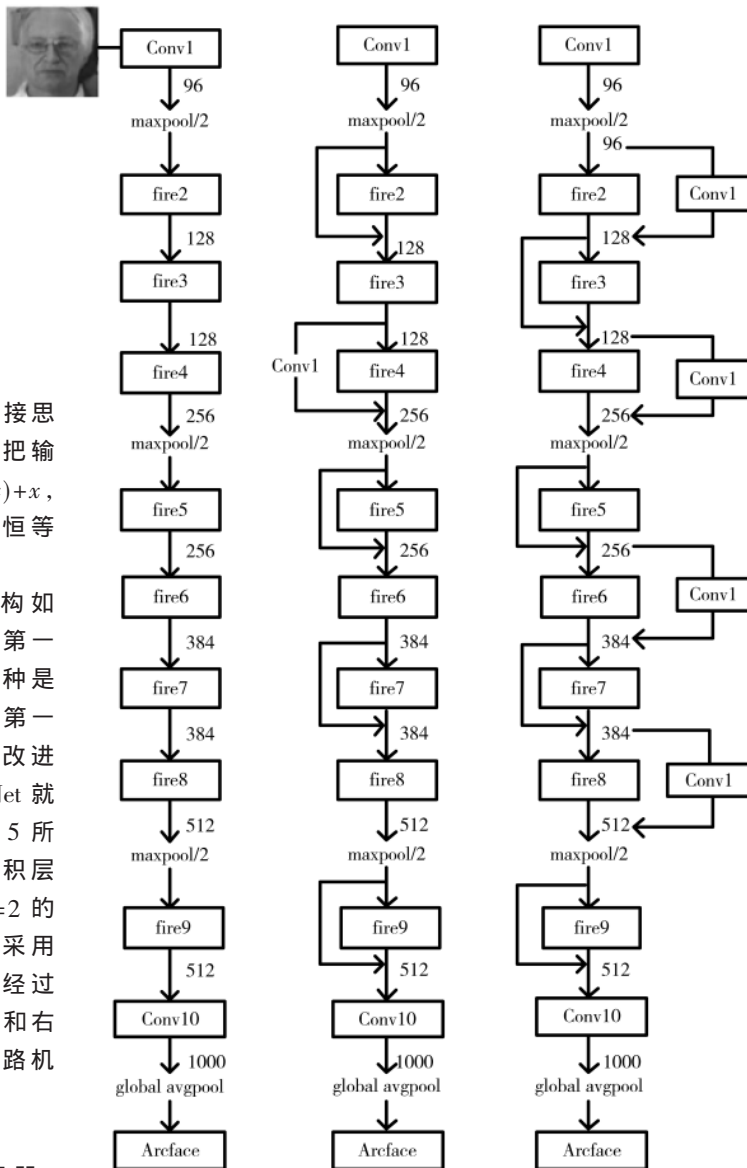


图5 基于 SqueezeNet 的不同网络架构



图6 预处理后的图片对比图

为 P_{same} , 所有的不同身份测试对为 P_{diff} , 阈值为 d , 测试对之间的欧氏距离之差为 $D(x_i, x_j)$, 那么正确接受次数 $TA(d)$ 、错误接受次数 $FA(d)$ 、正确拒绝次数 $TR(d)$ 以及错误拒绝次数 $FR(d)$ 分别为:

$$TA(d) = \{(i, j) \in P_{\text{same}}, \text{with } D(x_i, x_j) \leq d\} \quad (3)$$

$$FA(d) = \{(i, j) \in P_{\text{same}}, \text{with } D(x_i, x_j) > d\} \quad (4)$$

$$TR(d) = \{(i, j) \in P_{\text{diff}}, \text{with } D(x_i, x_j) > d\} \quad (5)$$

$$FR(d) = \{(i, j) \in P_{\text{diff}}, \text{with } D(x_i, x_j) \leq d\} \quad (6)$$

网络模型的识别率 $VAL(d)$ 为:

$$VAL(d) = \frac{|TA(d)| + |TR(d)|}{|P_{\text{same}}| + |P_{\text{diff}}|} \quad (7)$$

3.3 实验结果

为了能够将人脸识运用在边缘计算的硬件上(如 FPGA), 本文主要采用了 SqueezeNet 网络来进行人脸识别, 主要在 IJB-A 数据集和 lfw 数据集上进行了测试。在 IJB-A 和 LFW 数据集上测试了原始的 SqueezeNet 测试结果和加入了人脸归一化(FNM)预处理后人脸识别的结果, 如表 1、表 2 所示。可以看出, 在 IJB-A 数据集上对进行了 FNM 预处理的人脸识别的识别率有明显的提升, 由于 LFW 数据集都是正常的人脸, 所以识别率提高不是很多。Rank1 准确度就是某个类别的概率最高, 就认为预测结果属于哪种类别; Rank5 准确度选取 5 个最大概率的类别, 只要这 5 个类别中的一个和真实标签相同, 该预测结果就为真。

表 1 在 IJB-A 数据集上的测试结果 (%)

网络模型	Rank1	Rank5
SqueezeNet	74.3	83.8
SqueezeNet+FNM	83.2	85.5

表 2 在 LFW 数据集上的测试结果 (%)

网络模型	Rank1	Rank5
SqueezeNet	90.2	95.4
SqueezeNet+FNM	91.4	96.2

对模型进行了改进, 将损失函数改为 Arcface 损失函数, 并且将修改了损失函数的网络记为 MySqueezeNet, 添加残差网络后的识别率为如表 3 所示, 主要在 IJB-A 数据集上进行测试。通过比较, 修改了损失函数的网络比

表 3 不同网络模型的识别率

网络模型	Rank1/%	Rank5/%	模型大小/MB
SqueezeNet	74.3	83.8	4.8
MySqueezeNet	77.1	85.9	4.8
MySqueezeNet+simple ResNet	79.2	87.5	4.8
MySqueezeNet+complex ResNet	78.9	88	7.7
MySqueezeNet+simple ResNet+FNM	80.1	88.4	4.8

开始的网络 SqueezeNet 识别率提高了 2% 左右。并且对比修改的两种网络, 识别率虽然都得到了提高, 但是 simple+ResNet 的网络模型的大小是保持不变的, 因此计算量不会增加, 但加入了 complex+ResNet 网络以后, 明显增加了该网络的计算量。因此, MySqueezeNet+simple ResNet 网络更好。为了说明加入 FNM 是有用的, 因此在该网络上进行了测试, 发现在该网络的基础上加入 FNM 后识别率也得到了提高。因此, 本文增加的修改了损失函数、预处理方法和加入了残差网络模块都是有用的。综合识别率和消耗的资源考虑, 加入了 simple+ResNet 结构更好。

4 结论

在边缘计算中, 采用大型的网络会造成识别时间长、功耗高等问题, 因此本文采用了 SqueezeNet 网络来进行人脸识别。由于直接对网络进行人脸识别, 识别率可能不高。因此, 本文采用了预处理来对识别的图片先进行处理, 发现识别率得到了相应的提高; 同时, 改进了 SqueezeNet 网络的损失函数, 以及在网络中增加了 ResNet 模块, 这些处理方式都使得识别率得到了相应的提高, 最后通过测试表明了该方法是有效的。本文方法还有很大的提升空间, 后续将进一步增强网络的鲁棒性, 使得网络可以在当前经典数据集上的识别率更高。

参考文献

- [1] QIAN Y, DENG W, HU J. Unsupervised face normalization with extreme pose and expression in the wild[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019: 9843-9850.
- [2] MASI I, RAWLS S, MEDIONI G, et al. Pose-aware face recognition in the wild[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016: 4838-4846.
- [3] HU Y, WU X, YU B, et al. Pose-guided photorealistic face rotation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018: 8398-8406.
- [4] LIU X, VIJAYA KUMAR B V K, JIA P, et al. Hard negative generation for identity-disentangled facial expression recognition[J]. Pattern Recognition, 2018, 88: 1-12.
- [5] CAO J, HU Y, ZHANG H, et al. Towards high fidelity face frontalization in the wild[J]. International Journal of Computer Vision, 2020, 128(5): 1485-1504.
- [6] SAHA A, PRADHAN S N. Facial expression recognition based on eigenspaces and principle component analysis[J]. International Journal of Computational Vision & Robotics, 2018, 8(2): 190-200.
- [7] 孙雨浩, 陶洋, 胡昊. 基于低秩矩阵恢复和 Gabor 特征的遮挡人脸识别[J]. 微电子学与计算机, 2020, 37(3): 46-52.

- [8] XIANG J, ZHU G. Joint face detection and facial expression recognition with MTCNN[C]. IEEE Computer Society, 2017: 424-427.
- [9] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: a unified embedding for face recognition and clustering[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015: 815-823.
- [10] IANDOLA F N, Han Song, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[J]. arXiv preprint arXiv: 1602.07360, 2016.
- [11] GAIKWAD A S, EL-SHARKAWY M. Pruning convolution neural network (squeezenet) using taylor expansion-based criterion[C]. 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Louisville, KY, USA, 2018: 1-5.
- [12] DENG J, ZAFERIRIOU S. ArcFace for disguised face recognition[C]. 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Korea (South), 2019: 485-493.
- [13] JANG E, GU S, POOLE B. Categorical reparameterization with gumbel-softmax[J]. arXiv Preprint arXiv: 1611.01144, 2016.
- [14] 周光朕, 杜姗姗, 冯瑞, 等. 基于残差量化卷积神经网络的人脸识别方法[J]. 计算机系统应用, 2018, 27(8): 39-45.
- [15] 朱红, 陈清华, 刘国岁. 一种高速神经网络 HS-K-WTA 的研究[J]. 电子学报, 2002, 30(7): 1020-1022.

(收稿日期: 2020-09-30)

作者简介:

况朝青(1996-), 通信作者, 男, 硕士研究生, 主要研究方向: 深度学习与计算机视觉, E-mail: 2315650756@qq.com.

贺超(1990-), 男, 博士研究生, 主要研究方向: 光纤无线通信网络。

王均成(1996-), 男, 硕士研究生, 主要研究方向: 深度学习与计算机视觉。



扫码下载电子文档

(上接第 29 页)

- [14] DUVENAUD D K, MACLAURIN D, IPARRAGUIRRE J, et al. Convolutional networks on graphs for learning molecular fingerprints[J]. arXiv: 1509.09292, 2015.
- [15] REX Y. Graph convolutional neural networks for web-scale recommender systems[C]. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018.

(收稿日期: 2020-10-06)

作者简介:

张杰俊(1971-), 男, 硕士, 主要研究方向: 电信网络 ICT、云计算、人工智能技术。

唐颖淳(1975-), 女, 硕士, 主要研究方向: 电信领域人工智能技术。

季述卿(1996-), 男, 硕士, 主要研究方向: 机器学习与图神经网络。



扫码下载电子文档

“FPGA 及人工智能”专栏征稿启事

自 1984 年诞生以来, FPGA (Field Programmable Gate Array, 现场可编程逻辑门阵列) 因其优异的可定制性和可重配置特点得到了工业界和学术界的密切关注和深入研究, 并在诸多领域得到广泛应用。近年来, 随着人工智能、大数据的迅速发展, FPGA 在人工智能等领域的应用也受到业界的大力关注。可编程性、高能效比等特点使 FPGA 在人工智能领域的应用中展现出独特优势。为了促进 FPGA 在人工智能、大数据、边缘计算等新兴应用领域的应用研究和技术推广, 推动 FPGA 及人工智能领域的发展, 《电子技术应用》杂志拟于 2021 年第 12 期 (12 月 6 日出刊) 推出“FPGA 及人工智能”主题专栏。现面向相关领域专家学者征集相关稿件。欢迎新老读者大力关注, 踊跃投稿!

1. 稿件主题: 稿件内容包括但不限于以下主题:

- (1) 基于 FPGA 的人工智能研究, 如图像和语音处理、深度学习、机器学习、虚拟现实、神经网络与智能计算、基于大数据的人工智能技术。
- (2) FPGA 在其他领域 (工业、通信、医疗等) 的应用。
- (3) 人工智能领域相关的算法研究及硬件实现。

2. 稿件要求: 文章需具有创新性且未在其他期刊公开发表过。文中图表需清晰, 文字规范。详见《电子技术应用》投稿须知 (<http://www.chinaaet.com/paper/notice/>)。

3. 截稿日期: 2021 年 10 月 20 日。

4. 投稿方式: 请登录《电子技术应用》官网 (<http://www.chinaaet.com/>), 投稿页面中选择“FPGA 及人工智能”专栏投稿, 按要求提交。

专栏特约主编: 韩德强 北京工业大学 高级工程师

专栏编辑: 毕晓东 (010-82306085; bixd@chinaaet.com)

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所