

多通道卷积神经网络图像识别方法

易超人, 邓燕妮

(武汉理工大学 自动化学院 湖北 武汉 430070)

摘要:为了更好地利用图像数据中隐含的特征信息,将多方向梯度信息作为边缘信息的基本表达,提出了一种基于图像梯度的多通道卷积神经网络图像识别方法。先将图像进行 Sobel 算子处理,得到水平方向、垂直方向及两个对角方向的 4 个梯度图像。然后,建立 4 个多层卷积神经网络,学习 4 个不同方向梯度图像的特征。再将 4 个不同方向的特征进行随机化特征融合,得到样本的特征后经过批标准化处理。最后,通过分类器得到分类结果。在数据库 Cifar-10 和 MNIST 上进行了验证,验证结果表明:本文提出的模型具有较好的泛化能力,相比单通道卷积神经网络,在两个数据库中识别错误率分别降低了 9.85% 和 0.38%。

关键词:卷积神经网络;多通道;梯度图像;随机化特征融合;分类

中图分类号: TP391.41

文献标志码: A

0 引言

特征提取是模式识别领域中的核心技术,也是其难点之一。传统基于特征的识别方法都是先定义一种特征,再根据预先定义的特征进行分类识别。深度学习作为机器学习的重要发展,能够逐层地自动学习合适的表示特征,因而在众多领域中得到了广泛应用。卷积神经网络(convolutional neural network, CNN)作为深度学习的三大重要模型之一,继承了深度学习技术自动提取特征的优点,并且通过权值共享大大减少了所需要训练的参数,使 CNN 能快速处理高维图像,还具备一定的平移不变性,因而在图像分类识别等领域取得了较好的效果。目前,卷积神经网络已经被应用到人脸特征点检测^[1]、突出物体检测^[2]、行人检测^[3]、人脸识别^[4]和车型识别^[5]等领域,均取得了较好的效果。

上述检测和识别方法都是采用图像的直接输入。图像最基本的特征是边缘,边缘是图像信息最集中的地方,采用梯度信息作为边缘信息的基本表达来进行图像处理,在某些方面能够更好地利用图像信息。文献[6]提出了一种基于图像梯度的单样本人脸识别方法,采用梯度方法进行人脸识别具有光照不变性等优点,利用 B 样条滤波减小噪声对图像梯度信息的干扰,经验证优于传统的滤波方法。文献[7]将图像的梯度幅值、相位和结构相似度相结合,提出了一种新的利用梯度信息的图像质量评价模型—梯度相似度(gradient similarity, GSIM)模型,以及基于该模型的图像质量评价算法,经验证新模型能真实反映失真图像的视觉感知质量,具有较高的评价可靠性。

利用梯度信息处理图像具有一定的优点,但是上述方法只利用了图像水平方向和垂直方向的梯度信息,不能充分利用图像信息。为了更好地利用数据中隐含的特征信息,本文用水平方向、垂直方向及两个对角方向的 4 个梯度图像作为边缘信息,将多方向梯度信息作为 CNN 的输入,提出了一种基于图像梯度的多通道卷积神经网络图像识别方法。将 4 个梯度方向的特征信息输入到多通道卷积神经网络,使提取的图像特征更加有效,并且在分类层前加入了批标准化算法,进一步降低了图像识别的错误率。

1 卷积神经网络

1.1 卷积层和下采样层

1980 年,文献[8]提出新认知机,首次引进了 CNN 的概念,成为首个深度学习模型。2003 年,文献[9]对 CNN 进行了归纳总结。CNN 是一个多层非全连接的神经网络,每层由多个二维平面组成,而每

基金项目: 国家“863”计划基金项目(2015AA015904)

作者简介: 易超人(1993-)男,湖北荆州人,硕士生;邓燕妮(1962-)女,湖南益阳人,教授,博士,硕士生导师,主要研究方向为鲁棒优化控制理论及应用、计算机控制与应用、过程智能控制与优化技术等。

收稿日期: 2016-08-03

个平面由多个独立神经元组成。CNN 包含两种特殊的网络结构:卷积层和下采样层。卷积层和下采样层可以有多层,一般 CNN 前几层为卷积层和下采样层的交替,CNN 的深度由此体现^[10]。CNN 通过卷积层和下采样层的相互配合来学习原始图像的特征,并通过经典的误差反向传播(back propagation, BP)算法来调整参数,完成权值的更新。

卷积是图像识别中常用的算法,是指输出图像中的每个像素都是由输入图像对应位置的小区域的像素通过加权所得,这个小区域叫做局部感受野,区域的权值叫做卷积核。输入图像通过卷积运算后再加上偏置项,通过激活函数得到其特征图。文献[9]给出了卷积层的形式:

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} * K_{ij}^l + b_j^l\right),$$

其中: l 为层数; X_j^l 为在卷积层 l 层的第 j 个特征图; M_j 为输入层的感受野; K 为卷积核; b 为偏置; f 为神经元的激活函数,本文的激活函数选用的是修正线性单元(rectified linear unit, ReLU)函数。ReLU 激活函数可以防止梯度消失,缓解过拟合现象,取得更好的效果^[11]。

下采样层也叫池化层,即将图像分成一小块一小块的区域,对每个区域计算出一个值,然后将计算出的值依次排列,输出新的图像。这个过程相当于模糊滤波,可增加对图像特征提取的鲁棒性。本文池化方法采用的是平均池化。

1.2 批标准化

在图像处理过程中,对数据进行预处理,例如白化,甚至是简单的减均值图像预处理操作都能提高识别率。批标准化(batch normalization, BN)^[12]的思想则是对于每组批数据,在网络的每一层中,对各个特征分别进行标准化处理。文献[12]给出了算法流程,对于神经元在一次批处理过程中的全部输入,先计算均值和方差,再进行标准化处理,并在误差反向传播的过程中学习该算法的参数。一般批标准化可运用在任何层,且运用在神经元激活函数之前,即输入到神经元的数据通过加权运算后再经过批标准化处理,经过激活函数进行运算。

本文在经过随机化的特征融合后,输入到分类器前的数据运用批标准化,可以使输入到分类器中的数据由集中变得分散,取得更好的识别效果。在一定程度上,可以防止过拟合。

2 多通道卷积神经网络

图1是本文提出的多通道卷积神经网络模型结构图。首先,将输入图像经过 Sobel 算子得到水平方向、垂直方向及两个对角方向的4个梯度图像。然后,分别输入4个多层卷积神经网络,学习4个不同方向梯度图像的特征。再将4个不同方向的特征进行随机化特征融合,特征融合后的信息具有更好的鲁棒性^[13],得到样本的特征后再经过批标准化处理。最后,通过分类器得到分类结果并输出。

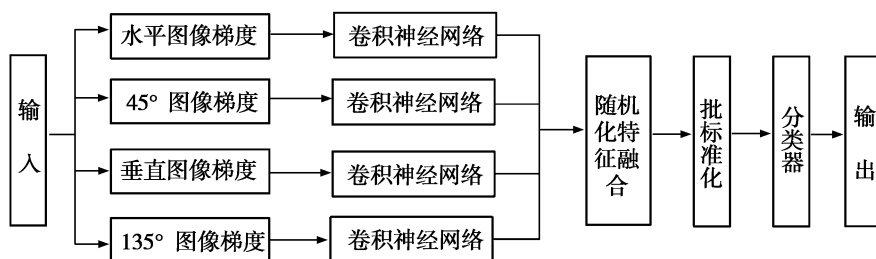


图1 多通道卷积神经网络模型结构图

对于一张 32×32 的图片,可得到其4个方向上的梯度信息,大小均为 32×32 。对于每个通道,图像梯度先经过8个 5×5 的卷积核得到8个 28×28 的特征图。然后,在池化大小为 2×2 下进行采样,得到8个 14×14 的特征图。再经过16个 5×5 卷积核后得到16个 10×10 特征图。最后,在池化大小为 2×2 下采样,得到16个 5×5 的特征图。每个网络都可以得到16个 5×5 的特征图,然后对4个通道的特征图进行随机化的特征融合。

随机化特征融合过程如图2所示。图2中: $A、B、C、D$ 分别代表不同通道的特征图,上标 j 表示每个通道的第 j 个特征图; $a、b、c、d$ 分别代表来自其对应特征图 $A、B、C、D$ 里的元素,下标数字代表对应特征

图元素的位置; M 代表随机化特征融合后的特征图。每次取 4 个网络中相同位置的 4 个特征图,再取 4 个特征图中相同位置的 4 个元素,然后用采样矩阵选取一个元素。采样矩阵按照其概率值大小随机选择,即元素值大的被选中的概率大。本文用 0-1 均匀分布来采样,将单位长度 1 按照 4 个元素值的大小分成 4 个区间,元素值越大,覆盖的区域越长,每个区间对应一个位置,然后随机生成一个 0~1 的数,这个数会落入某一区间。得到的采样矩阵为一个大小为 2×2 的矩阵,里面有一个元素为 1,其余元素为 0。通过这种随机化的特征融合,最后得到 16 个大小为 5×5 的特征图。把得到的特征图全连接处理后,得到一个大小为 400 的向量,经过批标准化处理后输入分类器进行分类。

3 实验结果及分析

选择普适物体数据库 Cifar-10 和手写数字库 MNIST,验证本文提出的基于梯度的多通道卷积神经网络模型。

3.1 实验方法

实验 1 是在数据库 Cifar-10 上进行验证。该数据库由 60 000 张 32×32 的 RGB 彩色图片构成,共分 10 类。其中,训练图片 50 000 张,测试图片 10 000 张,测试样本为交叉验证。本文将样本转化为灰度图像,归一化后在该灰度样本上进行了验证。实验中的深度自

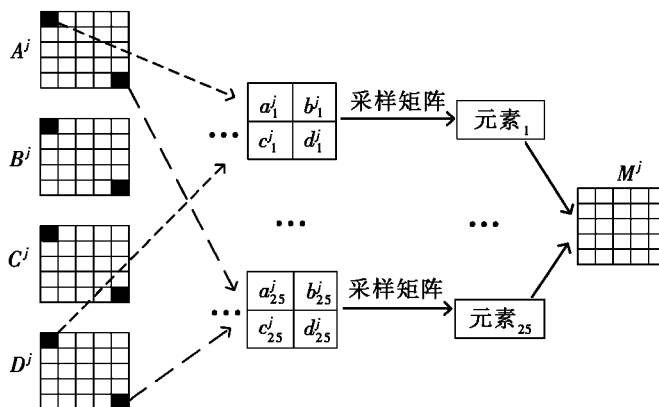


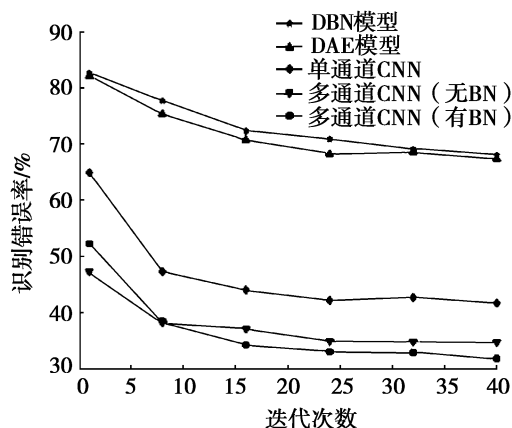
图2 随机化特征融合过程

动编码器 (deep auto-encoder, DAE) 和深度信念网络 (deep belief network, DBN) 模型的层数均为 [1 024 600 200 10]。单通道卷积神经网络包括 2 个卷积层和 2 个池化层,其中,第 1 个卷积层为 8 个卷积核,第 2 个卷积层为 16 个卷积核,大小均为 5×5 ,池化层大小均为 2×2 。单通道卷积神经网络输入的是进行归一化后的灰度图像,多通道卷积神经网络输入的是其灰度图像进行 Sobel 算子计算后的 4 个不同方向的图像梯度。多通道卷积神经网络的 4 个卷积神经网络模型均和单通道的模型相同,本文在多通道卷积神经网络模型中分类器前,对有无批标准化进行了验证对比。所有模型的分类器均为 softmax 分类器。

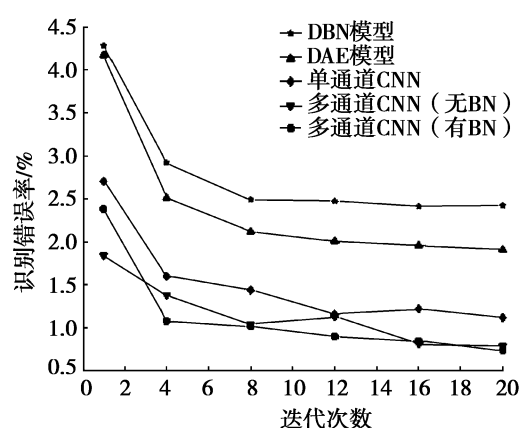
实验 2 是在数据库 MNIST 上进行验证。该数据库由 70 000 张 28×28 的灰度图片构成,共分 10 类,其中,训练图片 60 000 张,测试图片 10 000 张。同样将样本进行归一化处理。实验中的 DAE 模型和 DBN 模型的层数均为 [784 500 200 10]。单通道的卷积神经网络同样包括 2 个卷积层和 2 个池化层,两层卷积层卷积核的个数分别为 6 个和 12 个,大小均为 5×5 ,池化层大小均为 2×2 。实验内容与实验 1 基本类似。

3.2 结果分析

图 3 为 Cifar-10 和 MNIST 两种数据库在不同算法下的识别错误率。由图 3a 和图 3b 可以看出:在两种数据库上,本文提出的多通道 CNN (有 BN) 模型,比单通道 CNN 识别错误率分别降低了 9.85% 和



(a) Cifar-10 数据库在不同算法下的识别错误率



(b) MNIST 数据库在不同算法下的识别错误率

图3 Cifar-10 和 MNIST 数据库在不同算法下的识别错误率

0.38%。CNN 在图像识别方面比其他深度学习算法优越,这主要是因为 CNN 通过局部感受野和权值共享,能够有效地提取图像的特征,并具备一定的平移不变性。从图 3a 和图 3b 的单通道 CNN 和多通道 CNN(无 BN)的曲线可以看出:将图片的多方向梯度信息输入网络并做随机化的特征融合,比直接输入原始图片信息效果要好。这主要是因为将原始的数据直接输入卷积神经网络,虽然也能得到较低的识别错误率,但在模型结构较简单的条件下,不能够有效学习数据集中隐含的特征信息,而特征的有效性直接决定了分类性能。4 个梯度方向的特征信息能更全面利用数据中隐含的特征信息,两个数据库上取得的识别错误率更低,这也说明了将多方向梯度信息作为边缘信息的基本表达是有效的。本文将多通道模型分类层前有无批标准化算法进行了验证对比,运用批标准化后,Cifar-10 数据库经过 40 次迭代后识别错误率降低了 2.88%。由于 MNIST 数据库识别错误率较低,下降的幅度不是很明显,说明将集中的数据变得更分散,在一定程度上有利于防止过拟合,得到更好的效果。

4 结束语

(1) 卷积神经网络能自动提取图像的结构信息等特征,通过局部感受野、权值共享和下采样技术有效提取图像信息,并具备平移不变性,在图像识别上优于其他深度学习算法。

(2) 利用图像多方向梯度信息作为图像信息的表达,在某些方面可以更好地利用图像隐含的信息,并且通过随机化的特征信息融合,增加网络的鲁棒性,可以得到更好的识别效果。

(3) 批标准化使输入到分类器中的数据由集中变得分散,在一定程度上防止了过拟合,可以进一步降低识别错误率。

本文所提出的多通道卷积神经网络模型在数据库上可以取得更低的识别错误率,说明该方法是有效的,但此模型还较为复杂,多通道梯度卷积神经网络比单通道卷积神经网络的训练时间大约增加一倍。后续研究工作是优化模型的拓扑结构,减少误差反向传播时间,降低网络复杂度。

参考文献:

- [1] SUN Y, WANG X, TANG X. Deep convolutional network cascade for facial point detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland, Oregon, USA: IEEE, 2013: 3476–3483.
- [2] HE S F, LAU R W H, LIU W X, et al. SuperCNN: a superpixelwise convolutional neural network for salient object detection[J]. International journal of computer vision, 2015, 115(3): 330–344.
- [3] 苏松志, 李绍滋, 陈淑媛, 等. 行人检测技术综述[J]. 电子学报, 2012, 40(4): 814–820.
- [4] MOON H M, CHANG H S, PAN S B. A face recognition system based on convolution neural network using multiple distance face[J]. Soft computing, 2016, 20(2): 1–8.
- [5] DONG Z, PEI M T, HE Y, et al. Vehicle type classification using unsupervised convolutional neural network[C]// Proceedings of the 2014 22nd International Conference on Pattern Recognition (ICPR). Stockholm, Sweden: IEEE, 2014: 172–177.
- [6] 方蔚涛, 吴凡. B 样条和图像梯度的单样本人脸识别[J]. 重庆大学学报(自然科学版), 2013, 36(12): 111–116.
- [7] 马旭东, 闫利, 曹伟, 等. 一种新的利用梯度信息的图像质量评价模型[J]. 武汉大学学报(信息科学版), 2014, 39(12): 1412–1418.
- [8] FUKUSHIMA K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological cybernetics, 1980, 36(4): 193–202.
- [9] BEHNKE S. Hierarchical neural networks for image interpretation[M]. Berlin: Springer, 2003.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25(2): 1106–1114.
- [11] DAHL G E, SAINATH T N, HINTON G E. Improving deep neural networks for LVCSR using rectified linear units and dropout[C]// 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Vancouver, BC, Canada: IEEE, 2013: 8609–8613.
- [12] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]// Proceedings of the 32nd International Conference on Machine Learning. Lille, France: IEEE, 2015: 448–456.
- [13] 李欢欢, 万源, 董恒庆, 等. 分层纹理特征和梯度特征融合的图像分类[J]. 河南科技大学学报(自然科学版), 2015, 36(1): 52–57.

CHEN Liao^a, TAN Yubin^a, PAN Chaofeng^{a, b} (*a. School of Automotive & Traffic Engineering; b. Automotive Engineering Research Institute Jiangsu University Zhenjiang 212013, China*)

Abstract: To make the electric vehicle recover more energy during braking a new hydraulic braking actuating system of electric vehicle was designed. When electric cars were braked in low strength the mechanical friction braking force and motor regenerative braking force were decoupled. At this time, the braking force was completely provided by the motor regenerative braking force, which ensures the braking safety and efficient recovery of braking energy. The simulation model of the actuation system was established based on the advanced modeling environment for performing simulation of engineering systems (AMESim). According to the design scheme the physical experiment platform was set up. The simulation results and experimental results show that the braking actuating system is reasonable.

Key words: electric vehicle; hydraulic braking; AMESim software; experimental verification

CLC number: U463.52

Document code: A

Article ID: 1672-6871(2017)03-0025-05

Experiment on Reducing Emissions of Non-road Diesel Engine (30)

CHEN Zhanyao¹, JIA Yi¹, LI Ya'nan¹, MA Zhihao¹, ZHANG Heng² (*1. Vehicle & Transportation Engineering School Henan University of Science & Technology Luoyang 471003, China; 2. Maanshan Changlifa Machinery Manufacturing Alternative Translations Maanshan 243000, China*)

Abstract: Effects of injection pressure, injection timing, injector hole diameter and injector nozzle protrusion on exhaust emissions from a N490 naturally aspirated direct injection diesel engine were studied in order to reduce its exhaust emissions. The results show that increasing the injection pressure effectively reduces the exhaust smoke and specific fuel consumption of the diesel engine, but it will lead to the increase of NO_x emission. Increasing the injection timing effectively reduces the PM emission, but it will also increase NO_x emissions. Smaller nozzle number and diameter (5 × φ 0.21 mm) matched with higher pressure fuel injection pump effectively reduce NO_x and PM emissions. The suitable nozzle protrusion improves the emission performance of diesel engine. By optimizing and matching for the combustion system of diesel engine, NO_x + HC, CO, PM emissions of the diesel engine decrease to 6.23 g/(kW · h), 1.109 g/(kW · h), 0.522 g/(kW · h), respectively, which meets the China stage III emissions legislation of the non-road diesel engine.

Key words: non-road diesel engine; injection pressure; injection timing; injector hole diameter; nozzle protrusion

CLC number: TK421.5

Document code: A

Article ID: 1672-6871(2017)03-0030-05

Model and Algorithm of Rerouting under Adverse Weather with Moving and Shrinking (35)

WANG Lili, WANG Xin, WEI Fang (*Tianjin Key Lab of Operation Programming & Safety Technology of Air Traffic Management Civil Aviation University of China, Tianjin 300300, China*)

Abstract: To deal with constant-rate shrinking mobile danger zone under adverse weather, considering the effects of upper wind on flight tracks, a new systematic rerouting strategy including rerouting judging, two-stage rerouting trajectory inspection model and validation model was proposed. In the first two stages of dynamic rerouting, both the starting point of rerouting and the vector to ending were obtained by the designed heuristic algorithm. The trajectory inspection model was established to check potential conflicts between current course and dynamic convex polygonal danger zone. The conflicts were eliminated via the slightest adjustments on rerouting tracks by repeating the heuristic algorithm. The validation model was constructed to shorten the revised route. A simulation on Guiyang-Changsha route proved the effectiveness of the model and algorithm.

Key words: air traffic management; rerouting model; dangerous region with moving and shrinking; heuristic algorithm

CLC number: V355

Document code: A

Article ID: 1672-6871(2017)03-0035-06

• Electrical Engineering and Communication Automation and Computer •

Image Recognition Method of Multi-channel Convolutional Neural Network (41)

YI Chaoren ,DENG Yanni (*School of Automation ,Wuhan University of Technology ,Wuhan 430070 ,China*)

Abstract: To make better use of the feature information implied in the image database ,a multi-channel convolutional neural network image recognition method based on image gradients was proposed by making the multi-directional gradients information as basic expression of edge information. Firstly ,four gradient images of horizontal ,vertical and two diagonal directions were obtained by processing the image through Sobel operation. Then ,four multi-channel convolutional neural network was built to learn the features of gradient images of four different directions ,and the features of four different directions were fused by stochastic feature fusion. Finally , the classification results were obtained through the classifier after batch normalization process for the features of samples. The proposed model was verified on database Cifar-10 and MNIST. The results show that the proposed model has better generalization ability ,and its recognition error rates are reduced by 9.85% and 0.38% respectively compared to single-channel convolution neural network in the two databases.

Key words: convolutional neural network; multi-channel; gradient image; stochastic feature fusion; classification

CLC number: TP391.41

Document code: A

Article ID: 1672-6871(2017) 03-0041-04

New Type Z-source Inverter (45)

CHEN Xiao ,SU Hongsheng (*School of Automation & Electrical Engineering ,Lanzhou Jiaotong University , Lanzhou 730070 ,China*)

Abstract: In order to improve the output voltage of inverter direct current side ,a new type of Z-source inverter was proposed. In this topology ,the traditional inductance of Z-source in the network was replaced by the switch inductance ,and two switch inductor type quasi Z source inverter were connected in parallel to form a voltage boosting module. The operation principle of the circuit was discussed and verified by MATLAB simulation. The results of MATLAB simulation show that this topology is more powerful in the short run time and the voltage stress of a capacitor in the Z source network is effectively reduced.

Key words: new type Z-source inverter; topology; boost capability; capacitor voltage stress

CLC number: TM464

Document code: A

Article ID: 1672-6871(2017) 03-0045-04

Intrusion Protection System with Cognition Function (49)

CHEN Wei^a ,WU Huiqun^a ,TANG Lemin^a ,WANG Wei^b (*a. Medical School; b. Affiliated Hospital; Nantong University ,Nantong 226001 ,China*)

Abstract: An intrusion prevention system (IPS) combined with Soar was designed. By the cognitive technology of Soar ,IPS had the function of cognition. The simulation results show that Soar can make a comprehensive study on the source data package. Once Soar meets the packages with similar features ,these packets are directly labelled as scan packages ,and stored into the knowledge base which can be used as the evaluation standard later. The system achieves the purpose of self-defense. To a certain extent ,the demand of computer network security is met.

Key words: intrusion prevention system; security; cognitive; simulation

CLC number: TP393.081

Document code: A

Article ID: 1672-6871(2017) 03-0049-05

• Architecture •

Size Effect of Glass Fiber Reinforced Plastic Confined Concrete Column (54)

YU Yang ,YANG Yong ,ZHANG Yunfeng ,CHEN Sitong (*College of Civil Engineering & Architecture , Northeast Petroleum University ,Daqing 163318 ,China*)

Abstract: Three glass fiber reinforced plastic (GFRP) confined reinforced concrete columns and three unconfined reinforced concrete columns were tested and the finite element analysis was carried out on 15 specimens. The influence of different volume allocation rates on the mechanical properties of GFRP-confined reinforced concrete columns under axial compression was investigated. The results show that with the increasing of the size ,the ultimate compressive strength and the ultimate strain become smaller. At the same time ,the