

匹配实践要素与方案并提升精准度；虚拟场景构建模块协同VR/AR技术生成虚拟场景，支持调整方案，提升优化效率，强化适配。

4.2 形成课程融合育人格局

作为“三维融合”框架落地方式，以高职教育人才培养为遵循，聚焦基础、核心、拓展课程中党建元素分层融入，构建“基础课渗透-核心课强化-拓展课延伸”阶梯式育人格局，实现专业与思政教育同行。基础课程引入生成式人工智能辅助工具，探索党建元素符号化、具象化表达，各专业基础课程利用智能素材生成平台等指导学生创建资源，培养诠释党建内涵基础能力，如机械专业生成关联图解，让党建元素渗透基础知识学习。核心课程运用该技术，强化党建精神内涵转化，以真实项目为载体完成全流程实践，如电商专业分析关联点、生成建议，引导学生转化精神为专业实践，强化党建内涵与职业场景融合。拓展课程延伸应用场域，推动技术与思政教育创新融合，聚焦特色项目，发挥生成式人工智能与VR/AR技术协同优势，如建筑专业生成修复方案、构建虚拟场景，学生锻炼能力、深化认知，拓宽党建引领在职业教育的表现与应用维度。

4.3 打造“双师协同”团队

构建“双师型”教师团队，集党建素养、专业能力与技术水平于一体，通过“内培外引、校企协同”提升综合能力。党建导师由基层党组织书记担任，把握党建内涵与教育方向，结合党建工作为教学与实践提供资源和契机，指导学生实践，培育党性与责任感；专业教师由骨干教师组成，是团队中坚，深耕学科，融合党建内涵与专业内容，设计教学方案与实践项目，引导学生诠释党建精神，开展融合教学研究；技术专员由企业人工智能专家担任，提供技术支持与培训，熟悉前沿技术，为团队设计方案、攻坚难题，支撑创新应用，开展实操培训，推动技术与教学对接，服务党建引领与专业育人目标。

4.4 “党建教育+专业实践”为载体的融合发展实践

在理论构建基础上，我们选取两个代表性实践项目为突破口，以“党员带头、示范引领、师生共创”思路，探索生成式人工智能赋能党建与高职教育融合的实施路径。

4.4.1 党建教育创新：虚拟仿真智慧党建工作室建设

整合多学科力量，由党员教师骨干牵头建工作室，打造综合性党建育人平台，发挥党员教师示范引领作用，形成“党建带专业、党员带师生”机制。依托VR/AR技术和生成式人工智能，打造沉浸式党员教育创新阵地，破解传统党建教育问题，实现红色基因传承与专业能力培养统一，推动党建教育升级。核心功能模块有虚拟展馆、“党员讲党史”沉浸式党课，运行实行“预约管理+内容动态迭代”，由

党建导师与专业教师审核，学生党组织负责运营维护，党员突击队保障内容更新。以党员教师带队为驱动，联动示范岗与突击队实现双重突破，打破时空壁垒，借助人工智能挖掘数据，形成闭环创新链条。

4.4.2 专业实践融合：数字融合创新实训平台

为深化“党建+专业”融合育人成效，构建数字融合创新实训平台延伸工作室实践功能，将党建元素嵌入专业实践全过程，打造虚实结合的育人载体，以行业核心数字化技术构建专业实践模型，结合VR/AR技术与生成式人工智能，将党建元素融入职业教育实践教学体系，打造“思政内涵铸魂、专业技能立身、智能技术赋能”的实训平台，实现价值引领与能力培养协同增效。其核心功能模块包括红色地标数字化建模与VR漫游、“党建+”主题虚拟实践项目，将平台实践项目纳入课程体系并与学分挂钩，对接地方真实需求，构建“教学实训—项目研发—社会服务”闭环体系，推动党建元素从“文字表述”转化为“专业实践语言”，让学生在实践中提升技能、深化理解，实现“做中学思政、练技能悟初心”的育人效果。

两类实践载体功能互补，虚拟仿真智慧党建工作室侧重专业认知启蒙，数字融合实践平台侧重党建内涵深化，共同构建完整育人链条，为党建与高职教育融合提供实践范式。通过协同运行，形成融合发展模式，增强党建教育吸引力与实效性，推动专业实践创新，培育学生职业素养，实践积累的资源库为同类院校提供借鉴。

5 展望

在现代职业教育改革发展的时代背景下，以党建为核心引领、生成式人工智能为关键赋能、职业教育为实践载体的创新实践，是深入落实立德树人根本任务、加速推进职业教育数字化转型的核心举措。三者相互支撑、协同发力，构建起具有鲜明时代特色的高职专业建设新模式，有力推动高职教育高质量发展，为培养德技并修、又红又专的新时代高素质技术技能人才筑牢根基。

参考文献

- [1] 习近平.高举中国特色社会主义伟大旗帜为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告[M].北京:人民出版社,2022.
- [2] 教育部.《职业教育信息化融合发展行动计划》[Z].2022.
- [3] 高国希.课程思政的价值意蕴与生成路径[J].思想理论教育,2017(09):10-14.
- [4] 李华.课程思政与职业教育融合的实践路径探析[J].中国职业技术教育,2021(18):36-41.
- [5] 王树荫.坚持思政课建设与党的创新理论武装同步推进[J].思想教育研究,2020(03):12-16.

Application research of deep learning in image recognition

Guixin Xing Qian Meng

Zhongyuan University of Technology, Zhengzhou, Henan, 451191, China

Abstract

With the rapid advancement of artificial intelligence, deep learning has emerged as the core methodology for image recognition. By employing nonlinear mapping through multi-layer neural networks, it achieves automatic feature extraction and hierarchical representation, significantly enhancing recognition accuracy and generalization capabilities. Unlike traditional methods reliant on manual feature engineering, deep learning enables end-to-end learning driven by big data. This paper systematically analyzes the principles, mainstream models, and typical applications of deep learning in image recognition, with a focus on exploring the mechanisms and optimization strategies of CNN, RNN, and GAN in object detection, facial recognition, and building energy efficiency analysis. The research demonstrates that deep learning accelerates the intelligent transformation of image recognition, showcasing promising applications in autonomous driving, architectural monitoring, and security surveillance, while providing robust support for the development of intelligent vision systems.

Keywords

deep learning; image recognition; convolutional neural network; feature extraction

深度学习在图像识别中的应用研究

邢桂欣 孟茜

中原工学院, 中国·河南 郑州 451191

摘要

随着人工智能的快速发展,深度学习已成为图像识别的核心方法。其通过多层神经网络的非线性映射,实现特征的自动提取与层次表达,大幅提升识别精度与泛化能力。相较于依赖人工特征的传统方法,深度学习能够在大数据驱动下实现端到端学习。本文系统分析深度学习在图像识别中的原理、主流模型与典型应用,重点探讨CNN、RNN与GAN在目标检测、人脸识别和建筑节能分析中的机制与优化策略。研究表明,深度学习推动了图像识别的智能化进程,在自动驾驶、建筑与安全监控等领域展现出广阔前景,为智能视觉体系的发展提供了有力支撑。

关键词

深度学习; 图像识别; 卷积神经网络; 特征提取

1 引言

图像识别作为人工智能领域的重要方向,旨在让计算机具备理解与分析视觉信息的能力。随着计算资源和数据规模的提升,深度学习推动了图像识别技术从人工特征提取向自动特征学习的跨越。卷积神经网络(CNN)及其衍生模型在目标检测、人脸识别、建筑节能分析等领域取得显著突破。近年来,Transformer和Vision Transformer(ViT)等新架构的出现,进一步拓展了图像识别的深度与广度。然而,模型复杂度与计算成本的提升也带来了新的挑战。本文系统探讨深度学习在图像识别中的核心原理、关键技术与未来趋势,为智能视觉系统的优化与应用提供理论参考与实践启示。

【作者简介】邢桂欣(1989-),女,中国河南新乡人,硕士,讲师,从事人工智能研究。

2 深度学习与图像识别的理论基础

2.1 深度学习的基本原理

深度学习是机器学习的重要分支,其核心思想是通过多层神经网络构建非线性映射模型,实现数据特征的层次化表示。典型网络结构包括输入层、多个隐藏层与输出层,通过梯度下降与反向传播算法不断优化权重参数,使模型能够在高维数据空间中自动提取特征。深度学习在图像识别中的关键优势在于:无需人工设计特征、能够捕捉复杂的空间与语义关系、具备强大的泛化能力。随着GPU计算能力的提升与大规模标注数据集(如ImageNet)的出现,深度学习的训练效率与识别精度得到了显著提升,推动了计算机视觉的快速发展。

2.2 图像识别的主要任务与流程

图像识别的核心任务包括分类、检测、分割与识别等。其基本流程可分为数据预处理、特征提取、模型训练与结果预测四个阶段。在深度学习框架下,模型能够通过自动学习

提取边缘、纹理、形状及语义特征，实现对目标的精确分类。与传统方法相比，深度学习模型能够在端到端的训练过程中自适应优化特征空间，减少人工干预，提高模型的通用性与鲁棒性。

2.3 深度学习在图像识别中的优势

深度学习通过多层网络结构实现从低层到高层的特征抽象，使模型具备强大的非线性表达能力。其在大样本数据下表现出更高的精度与稳定性，能够有效应对复杂环境下的图像识别任务。此外，深度学习模型可结合迁移学习与预训练技术，实现小样本学习与跨领域适应，为图像识别在多领域应用奠定了基础。

3 卷积神经网络 (CNN) 在图像识别中的应用

3.1 CNN 的结构与工作机制

卷积神经网络 (CNN) 以“局部连接—权值共享—层次表征”为核心机制。卷积层利用卷积核在空间上滑动，提取边缘、纹理与形状等局部不变特征；通过步长与填充调控感受野与特征图尺寸，在保证空间分辨率与计算量间取得平衡。池化层（最大/平均池化）进行下采样，抑制噪声并增强平移不变性，同时降低参数规模。非线性激活提升网络对复杂映射的表达能力；批归一化 (BN) 稳定梯度分布、加速收敛；Dropout 在训练期随机失活神经元以抑制过拟合。在高层，卷积—归一化—激活的堆叠逐级聚合语义，配合全连接层或全局平均池化 (GAP) 完成类别判别。现代 CNN 还常引入残差/密集连接、注意力模块与多尺度特征金字塔，使网络既能保持梯度可传递性，又能兼顾细节与全局语义，实现对复杂场景的稳健表征。

3.2 典型 CNN 模型与性能比较

自 LeNet 奠定雏形后，AlexNet 凭借更深的结构、ReLU 与 GPU 训练在 ImageNet 上实现突破；VGG 以规则化的 3×3 小卷积深堆叠获得更强的特征表达，但参数量与计算代价较高。GoogLeNet 通过 Inception 多分支结构实现卷积核尺度自适应，有效压缩计算；ResNet 以残差连接显著缓解退化与梯度消失，使得上百层网络可稳定训练，成为后续模型基线。DenseNet 通过特征复用提升参数效率与信息流通；MobileNet 与 ShuffleNet 面向移动端，以深度可分离/组卷积降低 FLOPs；EfficientNet 采用复合缩放策略在网络深度、宽度与分辨率间协同优化，达到精度—效率更优平衡。综合比较，ResNet/ConvNeXt 类在通用性与可迁移性上表现稳定，EfficientNet 在算力受限场景具优势，DenseNet 在小数据与需强特征复用任务中更具性价比。模型选择应结合数据规模、延迟预算与部署平台协同权衡。

3.3 CNN 在典型应用领域的表现

目标检测方面，Faster R-CNN 以两阶段“候选框+分类回归”实现高精度，适合精细场景；YOLO、SSD 等一阶段方法以密集先验与多尺度特征图实现端到端推理，兼顾 mAP 与实时性，在视频监控与边缘设备部署中应用广泛。

人脸识别领域，FaceNet 以度量学习产生低维嵌入，ArcFace 通过角度间隔损失拉大类间距离、压缩类内方差，在姿态、光照与遮挡变化下保持鲁棒；配合关键点对齐与活体检测，已形成从采集到验证的工程闭环。总体来看，CNN 依托多尺度特征融合与结构正则化，在精度、稳定性与部署可达性上形成良好折中，并通过与注意力、Transformer 模块的协同，不断拓展在自动驾驶、安防与监测预警等关键行业的应用边界。

4 深度学习的扩展模型与创新应用

4.1 循环神经网络 (RNN) 与时序图像识别

循环神经网络 (RNN) 在图像识别中主要用于处理视频、动态场景及序列图像任务。其核心优势在于能捕捉时间维度上的连续变化特征，实现空间与时间信息的融合。标准 RNN 通过循环连接结构，将前一时刻的输出作为当前输入的一部分，从而建立序列间的依赖关系。然而，传统 RNN 在长序列训练中容易出现梯度消失或爆炸问题，限制了其对长期依赖的学习能力。长短期记忆网络 (LSTM) 与门控循环单元 (GRU) 通过引入输入门、遗忘门与输出门机制，有效控制信息流动，使模型能够保留关键时序信息。在行为识别、视频监控、交通流预测及动作识别等任务中，RNN 能通过时间序列分析理解目标运动轨迹与状态变化。结合卷积网络提取的空间特征，形成 CNN-RNN 混合模型，可进一步提升视频中动态对象的识别精度与场景理解能力，为智能监控与自动驾驶提供强有力的算法支撑。

4.2 生成对抗网络 (GAN) 在图像增强中的作用

生成对抗网络 (GAN) 通过生成器与判别器的博弈训练机制，实现了从噪声到高质量图像的映射学习。生成器负责合成接近真实分布的样本，判别器则用于区分真假样本，两者相互促进，使生成结果不断逼近真实图像。GAN 在图像修复、去噪、超分辨率重建、风格迁移及数据增强等领域表现突出。特别是在图像样本不足或质量不均的场景中，GAN 可有效扩充训练数据，改善深度模型的泛化能力。例如，在遥感影像与卫星数据处理中，GAN 可用于图像重建与分辨率提升，增强目标识别的精度与可用性。近年来，条件 GAN (cGAN) 与 CycleGAN 等改进模型进一步扩展了其在多域图像变换与无监督学习中的应用潜力，使图像识别系统在数据稀缺与噪声复杂的环境下依然具备稳定性能。

4.3 Transformer 模型与视觉识别新范式

Transformer 最初用于自然语言处理，其核心的自注意力机制可捕捉全局依赖关系，实现对输入信息的权重分配。近年来，该结构被成功引入视觉领域，形成 ViT 等模型。ViT 将图像划分为若干固定大小的 Patch 片段，并将其线性嵌入为特征序列，利用多头自注意力机制在全局范围内建模图像各区域间的关联性，从而突破卷积网络局部感受野的限制。相比传统 CNN，Transformer 在复杂场景、多尺度目标及长距离依赖建模中表现更优。后续的 Swin Transformer