

医学影像中关键点检测技术的应用与前景分析

陈黄菲*, 陈 胜#

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年1月13日; 录用日期: 2025年2月14日; 发布日期: 2025年2月26日

摘 要

随着深度学习技术的迅速发展, 关键点检测技术在医学影像分析中的应用受到广泛关注, 尤其在超声、CT和MRI等医学影像中表现出巨大的潜力。文章首先回顾了传统的关键点检测技术与基于深度学习的关键点检测技术在医学影像中的应用, 重点分析了卷积神经网络(CNN)、Hourglass网络和Transformer模型的特点与优势; 随后讨论了关键点检测在医学影像中的实际应用, 包括人体姿势估计、器官与肿瘤的分割与定位等领域的应用。此外, 文章还总结了当前技术面临的挑战, 如数据不足、图像噪声、跨设备泛化等问题, 并提出了可能的解决方案。最后, 结合深度学习的最新进展, 本文展望了医学影像中关键点检测技术的未来发展趋势, 旨在为医学影像分析中的关键点检测技术的研究与应用提供理论支持和发展思路。

关键词

医学影像, 关键点检测, 深度学习, 神经网络, 技术进展, 应用前景

Application and Prospect Analysis of Key Point Detection Technology in Medical Imaging

Huangfei Chen*, Sheng Chen#

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Jan. 13th, 2025; accepted: Feb. 14th, 2025; published: Feb. 26th, 2025

Abstract

With the rapid development of deep learning technology, the application of keypoint detection

*第一作者。

#通讯作者。

technology in medical image analysis has received widespread attention, especially in medical images such as ultrasound, CT, and MRI, showing great potential. The article first reviews the application of traditional keypoint detection techniques and deep learning based keypoint detection techniques in medical imaging, with a focus on analyzing the characteristics and advantages of convolutional neural networks (CNN), Hourglass networks, and Transformer models; Subsequently, the practical applications of keypoint detection in medical imaging were discussed, including human pose estimation, segmentation and localization of organs and tumors, and other fields. In addition, the article also summarizes the challenges currently faced by technology, such as severe data shortages, image noise, cross device generalization, and proposes possible solutions. Finally, based on the latest advances in deep learning, this article looks forward to the future development trends of keypoint detection technology in medical imaging, aiming to provide theoretical support and development ideas for the research and application of keypoint detection technology in medical image analysis.

Keywords

Medical Imaging, Key Points Detection, Deep Learning, Neural Networks, Technological Progress, Application Prospects

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

医学影像分析是现代医学中的重要研究领域,随着医学影像技术的不断进步,越来越多的影像数据被应用于疾病诊断、治疗规划和病例评估等领域[1]。在这些复杂的影像数据中,关键点检测技术[2]作为一种精准定位特定结构或区域的工具,在医学影像分析中扮演着至关重要的角色。关键点检测不仅仅限于简单的目标检测或位置标定[3][4],它还涉及对图像中关键结构的准确定位与分析,帮助医生更有效地进行诊断和治疗决策。

近年来,随着卷积神经网络(CNN)及其衍生模型[5]的迅速发展,医学影像中的关键点检测技术取得了显著的进展。传统的图像处理方法主要依赖于人工设计的特征[6],但图像模糊、变形等因素往往会影响处理的结果,难以处理复杂的医学影像数据。而深度学习技术凭借其自动学习特征的能力,在多种医学影像任务中表现出色,在关键点检测任务中也发挥了巨大的潜力。关键点检测技术在医学影像中的应用广泛,涵盖了器官定位[7]、姿势估计[8]、超声图像分析[9]等多个方面。例如,在超声影像中,关键点检测能够精准识别血管、心脏等关键部位的变化,对于早期诊断和疾病监测具有重要意义;在CT和MRI影像中,关键点检测可以辅助分割、器官结构分析等任务,为肿瘤的早期发现和治疗提供疾病支持。然而,尽管这一领域的研究取得了初步进展,仍然面临着数据严峻的困难、跨设备泛化能力差、模型计算资源要求较高的挑战。

本综述旨在回顾和总结医学影像中关键点检测技术的研究进展,尤其是深度学习方法的应用与发展。文章讨论了当前主流的关键点检测模型,如卷积神经网络、Hourglass网络、Transformer等模型,并展望了医学影像中关键点检测技术的未来发展方向。

2. 关键点检测的传统方法回顾

2.1. 原理及定义

医学图像中的关键点检测是通过识别和定位图像中特定点或区域,提取重要的空间信息,进一步辅

助临床诊断与治疗决策[10]。与其他计算机视觉领域中的关键点检测任务(如人脸、人体)姿态估计类似,医学影像的关键点检测同样要求精确定位和分析图像特征。然而,医学影像中的目标通常更加复杂和多样性,且对检测准确性的要求更高,这使得该领域的关键点检测面临更大的挑战。尽管如此,医学图像关键点检测的许多研究思路和技术灵感来源于这些技术领域,通过不断探索和改进相关算法和模型,推动了该技术的快速发展。接下来,本文将简要介绍医学图像关键点检测中应用的相关技术及发展趋势。

2.2. 边缘检测与角点检测

边缘检测和角点检测是传统图像处理技术中最基础、广泛应用的方法。在医学影像分析中,边缘检测常用于识别器官或组织的边界,从而辅助定位关键点的定位[11]。常见的边缘检测算法,如 Canny 边缘检测算子,通过检测图像中的亮度变化来提取边缘特征[12]。然而,传统的 Canny 算法在处理灰度不均匀、边缘模糊的 MRI 影像时,往往存在边缘检测精度不足和假边缘的问题。为解决这些问题,改进的 Canny 算法[13]通过图像增强、中值滤波去噪、加权方向模板和自适应阈值等优化措施,显著提高了 MRI 图像的边缘检测精度,减少了假边缘的产生,增强了算法的自适应性。

另一方面,角点检测算法(如 Harris 角点检测[14])用于识别图像中的关键角点,这些角点通常与医学影像中的重要解剖结构(如骨骼、器官边缘等)相关。虽然这些方法计算效率高且实现简单,但它们的性能对图像质量高度依赖,并且对噪声、亮度变化等因素非常敏感,这限制了其在复杂医学影像中的应用。因此,虽然传统方法在简单场景中有效,但它们在医学影像中的普适性和鲁棒性仍然面临着巨大的挑战。

2.3. 基于模板匹配的方法

模板匹配是一种用于检测图像中和给定模板区域的传统方法。通过与预定义的模板进行匹配,可以快速识别出感兴趣区域的关键点[15]。在医学影像分析中,模板匹配常用于定位器官、组织以及病变区域,尤其适用于结构相对规则的影像。该方法的基本原理是通过计算图像与模板之间的相似度,找到最佳匹配区域,通常使用相关性度量或互信息作为匹配标准。

为了提高模板匹配方法的鲁棒性,尤其在处理具有旋转、缩放以及姿势变化的图像时,研究者提出了改进的模板匹配技术。例如,尺度不变特征变换(SIFT)和加速稳健特征(SURF)算法通过提取图像中的局部特征点[16],在不同的尺度、旋转和变形下保持较高的匹配精度。这些方法不仅能够增强模板匹配的稳定性,还能在复杂环境中提取更为精准的特征,提高匹配效果。

2.4. 基于机器学习的特征分类方法

传统的机器学习方法,如支持向量机(SVM)、随机森林等方法[17][18],也被广泛应用于医学影像中的关键点检测。这些方法依赖于手工提取的特征,如形状、纹理等和边缘特征,通过训练分类器对图像中的关键点进行定位和标定在此过程中,特征提取需要涉及图像分割的步骤,包括边缘检测、区域分割和纹理分析等,从而为分类器提供有意义的信息。通过对不同的特征进行学习,机器学习模型能够识别图像中的关键点并进行精确的标定。

支持向量机(SVM)作为一种常见的分类器,通过构建超图表将不同类别的样本分开,进而实现关键点的定位。而随机森林通过集成多个决策树进行投票,最终确定目标区域的位置这些方法通过对手工特征的有效组合,能够在一定程度上克服传统图像处理方法的限制方法,尤其是在较复杂的医学影像任务中,能够提供较准确的结果。因此,基于机器学习的特征分类成为医学影像分析中重要的技术手段之一。

3. 深度学习阶段的发展梳理

深度学习中卷积神经网络(CNN)[19]在图像处理领域的突破性进展,推动了医学图像分析的发展。与

传统的图像处理方法相比, 深度学习技术的优势在于其强大的自动化训练过程和对图像特征的自动提取和处理能力[20]。通过学习, CNN 能够自动从大量标签数据中学习到有特征, 需要人工设计复杂的特征提取算法。这使得深度学习适用于医学图像中的关键点检测任务, 能够自动从原始图像中提取出复杂的空间信息, 定位图像中的关键点。

3.1. 卷积神经网络(CNN)在医学影像中的应用

卷积神经网络(CNN)作为深度学习领域的基础模型, 在医学图像分析中得到了广泛应用。CNN 通过多个卷积层[21]、池化层[22]和全连接层[23]提取图像中的特征, 有效捕捉图像的空间信息。最初, 主要用于图像分类任务, 但随着计算能力的提升和大规模数据集的构建, CNN 逐渐扩展到医学影像中更复杂的任务, 如关键点检测[24]、器官分割[25]等。

在关键点检测任务中, CNN 能够自动学习医学图像中的重要特征。基于 CNN 的关键点检测模型通常将图像中的关键点视为特征点, 通过对每个像素位置的预测来实现关键点的定位。例如, 在脑部 MRI 图像中, CNN 模型能够根据图像的不同结构, 准确识别并脑定位的边界或脑血管的关键节点。与传统的肿瘤边缘检测、模板匹配等技术相比, CNN 方法具有更强的鲁棒性和更高的准确性, 尤其在处理复杂性较高、噪声分布较广的医学图像时, 表现出明显的优势。

3.2. Hourglass 网络

沙漏网络(Hourglass Network)是一种高精度的深度学习架构。其核心原理是通过多个编码器和解码器模块来实现多像素的特征提取与预测[26]。通过编码器部分阶梯下采样将图像输入的空间信息压缩, 并提取的图像高层次信号特征; 解码器部分则通过阶梯上采样图像恢复的空间分辨率, 以便准确地预测关键点的位置。这种结构使得沙漏网络能够在不同层次上捕获图像中的细节, 尤其适用于处理复杂的姿态估计和关键点定位问题。

在医学影像中的关键点检测任务中, Hourglass 网络能够高效处理各种医学图像, 如器官定位等。通过网络的多层次特征学习, Hourglass 网络能够精准定位图像中的关键点位置, 减少了传统方法中手动特征提取和多步处理的复杂性。在处理不同的测量和结构的医学图像时, 沙漏网络表现增强了鲁棒性和准确性。

3.3. 热力图回归方法

热力图回归方法(Heatmap Regression) [27]是近年来在医学影像关键点检测中取得显著进展的一种深度学习方法。核心思想是通过回归模型将图像中的每个关键点定位映射到一个热力图中, 每个热力图的像素值代表该位置为某个关键点的概率。最终, 通过热力图的热点位置来确定关键点的坐标。该方法通过构建网络输出多个热力图, 每个热力图对应一个关键点, 通过最小化预测热力图与真实热力图之间的差异来优化模型。在眼底图像视网膜血管检测任务中[28], 通过构建网络输出多个热力图, 每个热力图对应一个关键点, 模型通过最小化预测热力图与真实热力图之间的差异来优化训练过程。该方法实现了优异的性能, 能够准确定位血管交叉点和分叉点的位置。与传统的图像分割或边界检测方法相比, 热力图回归具有更强的鲁棒性, 尤其是在处理较大的噪声方面、背景复杂的医学图像时, 能够保持上述的准确性。

3.4. Transformer 关键点检测

Transformer 模型[29]通过捕捉序列中各位置之间的长期依赖关系, 已在自然语言处理领域取得了突破性的成果。该模型的核心原理是利用多头自注意力机制(Multi-Head) [30], 它能够建模序列中不同位置

之间的相互关系, 并且通过学习不同位置之间的注意力权重, 对输入特征进行动态调整, 从而更准确地捕捉到全局信息。与传统的神经网络(CNN)在局部范围内的限制不同, Transformer 能够在处理复杂空间关系时, 基于全局上下文信息更好地理解图像中的各个要素之间的关联。由于其优越的全局建模能力, Transformer 模型已经包括多个领域, 图像处理、目标检测和姿势估计等方面, 表现出优异的性能。

受到 Transformer 在 NLP 领域应用的影响, 研究人员开始将其结构和思想引入医学图像分析领域。传统的图像处理方法往往难以成功处理图像信息中的复杂空间和长程依赖, 而 Transformer 的自注意力机制能够有效地解决这个问题。在医学影像中, 关键点检测通常需要考虑解剖结构、损伤区域和周围组织之间的复杂关系, 这使得全局信息的捕捉视野变得极为重要。因此, Transformer 能够在复杂医学图像中抓取出各关键点之间的长期依赖关系, 极大提升了关键点检测的准确性。

TransPose [31]是基于 Transformer 的一个关键点检测方法, 它通过内在的注意力机制揭示了关键点之间的空间依赖关系, 并在人体姿势估计任务中取得了良好的效果。TransPose 模型结合了 Transformer 和卷积神经网络的优势, 既利用 Transformer 增强了全局上下文建模能力, 又保留了 CNN 在局部特征提取方面的优势。通过这种结合, TransPose 能够更好地捕捉医学图像中各个关键点之间的空间关系, 尤其是在处理涉及多个解剖区域的复杂医学图像时, 这种优势极其突出。

随着 Transformer 模型在医学影像领域的深入研究, 越来越多的创新应用出现, 如 TransUNet [32], 引入 Transformer 作为特征编码器并保留 UNet 的良好架构, 在图像分割任务中取得优秀成果。Vision Transformer (ViT) [33]通过引入空间注意力机制, 能够增强局部区域特征的建模, 进一步提升图像中细粒度信息的捕捉能力。这些方法不断推动医学影像分析技术的发展, 并在解剖学、病理学、肿瘤学等医学领域的检测任务中, 取得了越来越显著的成果。展望未来, Transformer 模型将继续在医学影像分析中发挥重要作用, 成为提高关键点检测准确性和效率的核心技术之一。

4. 不足与展望

4.1. 不足

4.1.1. 数据问题与数据集的不足

数据质量和数据量是深度学习技术成功应用的关键因素。然而, 在医学影像领域, 数据集普遍存在样本量不足、数据质量不高等问题, 这限制了模型的泛化能力。在关键点检测任务中, 医学影像的标注通常由专业医生完成, 但标注工作繁琐且主观因素过大, 这可能导致数据集质量不高, 从而影响模型的性能。

此外, 许多医学影像数据集具有较高的隐私性, 这使得在构建大规模数据集时面临法律和伦理挑战。尽管近年来一些开源医学影像数据集(如 ISIC [34]、BraTS [35]、LUNA [36]等)逐渐被开放, 但相比其他领域的深度学习任务, 医学影像领域的公开数据集仍然惊人不足。

4.1.2. 计算效率与实时性问题

深度学习模型对计算资源的需求很高, 这在处理医学影像时极为突出, 尤其是在复杂的神经网络和其他大型模型的训练中, 通常需要大量的计算资源, 而在许多医疗资源有限的情况下, 这可能成为应用深度学习技术的主要障碍。同时, 医学影像的存储和处理负担也很重, 因此需要对模型进行优化, 以提高计算效率和处理速度, 从而保证在临床诊断中能够实时查看和分析图像数据。

4.1.3. 临床应用的解释性与可信性

深度学习模型的“黑箱”特性在临床应用中缺乏透明度。医生和临床专家往往难以理解模型做出决策的原因, 这可能影响他们对模型预测结果的信任。尤其是在医学影像分析中的关键点检测任务中, 医

生需要能够理解模型决策的决策过程, 才能更有效地利用其辅助诊断和治疗。因此, 提高模型的可解释性, 帮助医生理解和信任深度学习模型的决策过程, 成为其广泛应用于临床的关键挑战之一。

4.1.4. 临床验证与大规模应用的挑战

目前, 大多数深度学习模型仍处于研究阶段, 缺乏充分的临床验证和实际应用数据。这些模型可能在不同的医院、设备和患者群体中表现出差异, 从而限制了其广泛应用。为了使深度学习模型学习技术在临床实践中真正发挥作用, 必须通过多中心临床试验进行验证, 确保模型能够适应认知的医疗场景。

4.2. 展望

针对当前的问题, 深度学习在医学影像中关键点检测的应用仍具有广阔的发展前景。

针对数据问题, 未来可以通过数据增强、跨机构数据共享和合成数据生成方式有效解决数据稀缺问题。生成对抗网络(GAN [37])等技术可以用于生成高质量的医学影像数据, 补充标注样本同时, 推动数据共享平台的发展、强化数据隐私保护措施, 提升数据多样性和质量, 将是解决数据问题的关键。

为了提升模型的泛化能力, 未来可以通过迁移学习(Transfer Learning [38])、自监督学习(Self-supervised Learning [39])以及多任务学习(Multi-task Learning [40])等方法, 增强模型的鲁棒性。该方法能够帮助模型从小型药品数据中提取更多的普适特征, 从而提升其在不同医院、设备和患者群体中的表现, 减少过度和数据偏倚问题。

在计算效率方面, 未来的研究可以通过轻量化网络架构(如 MobileNet [41]、EfficientNet [42]等)、模型压缩(如剪枝、量化等)以及边缘计算等技术, 降低模型的计算和存储负担, 从而能够在资源有限的设备上高效运行, 这对于推动深度学习技术在临床中的广泛应用至关重要。

对于可解释性问题, 未来可以结合可解释性 AI (Explainable AI, XAI) [43]技术, 提高模型的透明度, 帮助医生更好地理解和信任深度学习模型的决策过程。通过可视化技术和可解释性方法, 医生能够清晰地看到模型如何根据影像数据进行关键点检测, 这将为临床决策提供强有力的支持。

最后, 临床验证仍然是深度学习模型广泛应用的关键。随着多中心临床试验的开展及大规模数据集的积累, 深度学习模型有望在多种医疗环境中得到验证和应用, 为临床医生提供更多精准的辅助诊断工具, 推动医学影像分析技术的发展, 并进一步高效优化诊疗流程中的关键点检测。

5. 总结

本文回顾并分析了医学影像中关键点检测技术的发展现状, 重点探讨了传统方法与深度学习方法在该领域的应用, 物质未来发展方向进行了展望。医疗和辅助诊断领域取得了显著的成果, 但仍面临数据质量与数量、模型泛化能力、计算效率和可解释性等方面的挑战。

在传统方法中, 基于人工特征提取和传统机器学习的关键点检测技术, 逐步被深度学习方法所取代。深度学习模型能够从大规模数据中自动学习特征, 减少了人工特征工程的依赖, 并显著提升了检测性能。然而, 由于医学影像的复杂性和数据集的限制, 深度学习方法在实际应用中仍面临数据稀缺、样本单一、以及计算资源要求等问题。因此, 如何解决这些瓶颈, 成为深度学习技术广泛应用的关键。

本文探讨了深度学习在医学影像中关键点检测的未来发展, 提出了一些潜在的研究方向。首先, 可以通过数据增强、迁移学习和自监督学习等技术来克服数据稀缺的问题; 其次, 模型压缩、量化以及边缘计算等手段可用于提升计算效率; 此外, 加强可解释性 AI 的研究, 提升模型的透明度和临床可用性, 也是未来研究的重要方向; 最后, 通过多中心临床试验验证模型的实际效果, 将有助于推动深度学习技术的广泛应用。

随着技术的不断进步和跨学科合作的深入, 深度学习在医学影像分析中的应用前景, 预计将进一步

推动精准医疗的发展, 提高诊断和治疗效率。在未来, 随着数据质量的提升、计算资源的优化、以及临床验证的加强, 深度学习技术必将成为医学影像分析工作中的核心工具, 助力医生进行更高效、精准地诊断, 最终实现个性化、精准的医疗服务。

基金项目

国家自然科学基金 81101116。

参考文献

- [1] 俞瑞雪. 医学影像技术在医学影像诊断中的临床应用分析[J]. 国际全科医学, 2023, 4(3): 97-99.
- [2] Shu, C., He, Y. and Sun, Q. (2017) Point Cloud Registration Based on Convolutional Neural Network. *Laser & Optoelectronics Progress*, **54**, Article ID: 031001. <https://doi.org/10.3788/lop54.031001>
- [3] 罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. 电子学报, 2020, 48(6): 1230.
- [4] 高涵, 张明路, 张小俊, 等. 机械臂绝对定位精度标定关键技术综述[J]. 计算机应用研究, 2017, 34(9): 2570-2576.
- [5] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [6] 余东行, 郭海涛, 张保明, 等. 级联卷积神经网络的遥感影像飞机目标检测[J]. 测绘学报, 2019, 48(8): 1046-1058.
- [7] 孙冠雄. 人体关键点检测模型及其向心脏关键点定位的迁移研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2018.
- [8] 李梦荷, 许宏吉, 石磊鑫, 等. 基于骨骼关键点检测的多人行为识别[J]. 计算机科学, 2021, 48(4): 138-143.
- [9] 杨恒, 顾晨亮, 胡厚民, 等. 嵌入卷积增强型 Transformer 的头影解剖关键点检测[J]. 中国图象图形学报, 2023, 28(11): 3590-3601.
- [10] 韩冬, 李其花, 蔡巍, 等. 人工智能在医学影像中的研究与应用[J]. 大数据, 2019, 5(1): 39-67.
- [11] 章为川, 孔祥楠, 宋文. 图像的角点检测研究综述[J]. 电子学报, 2015, 43(11): 2315-2321.
- [12] 陆兴娟, 吴震宇. 图像边缘检测算法研究[J]. 现代电子技术, 2010, 33(6): 128-130.
- [13] 李健, 刘孔宇, 任宪盛, 等. 基于自适应阈值的 Canny 算法在 MRI 边缘检测中的应用[J]. 吉林大学学报(工学版), 2021, 51(2): 712-719.
- [14] 张小洪, 李博, 杨丹. 一种新的 Harris 多尺度角点检测[J]. 电子与信息学报, 2007, 29(7): 1735-1738.
- [15] 田娟, 郑郁正. 模板匹配技术在图像识别中的应用[J]. 传感器与微系统, 2008, 27(1): 112-114.
- [16] 张志强, 施文华. 改进的尺度不变特征变换算法并行加速双目测距系统及其实现[J]. 激光与光电子学进展, 2019, 56(14): 201-210.
- [17] Chandra, M.A. and Bedi, S.S. (2018) Survey on SVM and Their Application in Image Classification. *International Journal of Information Technology*, **13**, 1-11. <https://doi.org/10.1007/s41870-017-0080-1>
- [18] Hu, J. and Szymczak, S. (2023) A Review on Longitudinal Data Analysis with Random Forest. *Briefings in Bioinformatics*, **24**, bbad002. <https://doi.org/10.1093/bib/bbad002>
- [19] Sutskever, I. (2014) Advances in Neural Information Processing Systems.
- [20] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90.
- [21] Gao, L., Chen, P. and Yu, S. (2016) Demonstration of Convolution Kernel Operation on Resistive Cross-Point Array. *IEEE Electron Device Letters*, **37**, 870-873. <https://doi.org/10.1109/led.2016.2573140>
- [22] Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S. and Miao, Y. (2021) Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sensing*, **13**, Article No. 4712. <https://doi.org/10.3390/rs13224712>
- [23] Nassif, A.B., Shahin, I., Attili, I., Azzeh, M. and Shaalan, K. (2019) Speech Recognition Using Deep Neural Networks: A Systematic Review. *IEEE Access*, **7**, 19143-19165. <https://doi.org/10.1109/access.2019.2896880>
- [24] Zheng, Y., Liu, D., Georgescu, B., Nguyen, H. and Comaniciu, D. (2015) 3D Deep Learning for Efficient and Robust Landmark Detection in Volumetric Data. *MICCAI 2015*, Munich, 5-9 October 2015, 565-572. https://doi.org/10.1007/978-3-319-24553-9_69
- [25] Hesamian, M.H., Jia, W., He, X. and Kennedy, P. (2019) Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *Journal of Digital Imaging*, **32**, 582-596. <https://doi.org/10.1007/s10278-019-00227-x>

- [26] Susanto, Y., Livingstone, A.G., Ng, B.C. and Cambria, E. (2020) The Hourglass Model Revisited. *IEEE Intelligent Systems*, **35**, 96-102. <https://doi.org/10.1109/mis.2020.2992799>
- [27] Luo, Z., Wang, Z., Huang, Y., Wang, L., Tan, T. and Zhou, E. (2021) Rethinking the Heatmap Regression for Bottom-Up Human Pose Estimation. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 19-25 June 2021, 13264-13273. <https://doi.org/10.1109/cvpr46437.2021.01306>
- [28] Hervella, Á.S., Rouco, J., Novo, J., Penedo, M.G. and Ortega, M. (2020) Deep Multi-Instance Heatmap Regression for the Detection of Retinal Vessel Crossings and Bifurcations in Eye Fundus Images. *Computer Methods and Programs in Biomedicine*, **186**, Article ID: 105201. <https://doi.org/10.1016/j.cmpb.2019.105201>
- [29] Han, K., Xiao, A., Wu, E., et al. (2021) Transformer in Transformer. *Proceedings of the 35th International Conference on Neural Information Processing Systems*, 6-14 December 2021, 15908-15919.
- [30] Shamshad, F., Khan, S., Zamir, S.W., Khan, M.H., Hayat, M., Khan, F.S., et al. (2023) Transformers in Medical Imaging: A Survey. *Medical Image Analysis*, **88**, Article ID: 102802. <https://doi.org/10.1016/j.media.2023.102802>
- [31] Yang, S., Quan, Z., Nie, M. and Yang, W. (2021) Transpose: Keypoint Localization via Transformer. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, 11-17 October 2021, 11802-11812. <https://doi.org/10.1109/iccv48922.2021.01159>
- [32] Lin, A., Chen, B., Xu, J., Zhang, Z., Lu, G. and Zhang, D. (2022) DS-TransUNet: Dual Swin Transformer U-Net for Medical Image Segmentation. *IEEE Transactions on Instrumentation and Measurement*, **71**, 1-15. <https://doi.org/10.1109/tim.2022.3178991>
- [33] Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., et al. (2023) A Survey on Vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 87-110. <https://doi.org/10.1109/tpami.2022.3152247>
- [34] Kassem, M.A., Hosny, K.M. and Fouad, M.M. (2020) Skin Lesions Classification into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning. *IEEE Access*, **8**, 114822-114832. <https://doi.org/10.1109/access.2020.3003890>
- [35] Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al. (2015) The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, **34**, 1993-2024. <https://doi.org/10.1109/tmi.2014.2377694>
- [36] Naseer, I., Akram, S., Masood, T., Jaffar, A., Khan, M.A. and Mosavi, A. (2022) Performance Analysis of State-of-the-Art CNN Architectures for Luna16. *Sensors*, **22**, Article No. 4426. <https://doi.org/10.3390/s22124426>
- [37] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B. and Bharath, A.A. (2018) Generative Adversarial Networks: An Overview. *IEEE Signal Processing Magazine*, **35**, 53-65. <https://doi.org/10.1109/msp.2017.2765202>
- [38] Pan, S.J. (2020) Transfer Learning. *Learning*, **21**, 1-2.
- [39] Jaiswal, A., Babu, A.R., Zadeh, M.Z., Banerjee, D. and Makedon, F. (2020) A Survey on Contrastive Self-Supervised Learning. *Technologies*, **9**, Article No. 2. <https://doi.org/10.3390/technologies9010002>
- [40] Zhang, Y. and Yang, Q. (2017) An Overview of Multi-Task Learning. *National Science Review*, **5**, 30-43. <https://doi.org/10.1093/nsr/nwx105>
- [41] Sinha, D. and El-Sharkawy, M. (2019) Thin MobileNet: An Enhanced MobileNet Architecture. 2019 *IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, 10-12 October 2019, 280-285. <https://doi.org/10.1109/uemcon47517.2019.8993089>
- [42] Koonce, B. (2021) EfficientNet. In: Koonce, B., Ed., *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*, Apress, 109-123. https://doi.org/10.1007/978-1-4842-6168-2_10
- [43] Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D. and Zhu, J. (2019) Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. In: Tang, J., et al., Eds., *Natural Language Processing and Chinese Computing*, Springer International Publishing, 563-574. https://doi.org/10.1007/978-3-030-32236-6_51