

基于GNN的代理辅助U-Net架构进化方法

吴润佳

广东工业大学数学与统计学院, 广东 广州

收稿日期: 2025年2月28日; 录用日期: 2025年3月27日; 发布日期: 2025年4月3日

摘要

视网膜血管切片在诊断和治疗疾病状况方面起着重要作用。基于识别神经网络的技术在这一领域取得了许多不同的突破, 但如何设计最佳的网络结构是一个具有挑战性的问题。神经结构搜索在现实世界的程序中面临着挑战, 例如太多的搜索空间和太长的学习时间。为了有效地解决这些问题, 该数据展示了基于先进算法的U-Net架构搜索方法。与添加网络参数不同, 这种方法可以提高网络性能。同时, 该架构结合了代理模型和进化算法, 以降低搜索过程中的计算成本。本文使用神经网络作为替代模型来增强变异算法, 并通过绘制共同对象来执行匹配措施。同时, 本文加速了遗传计算方法的收敛速度。实验结果表明, 算法在灵敏度和AUROC等关键指标上仍优于最先进的方法。

关键词

代理模型, 进化算法, 神经架构搜索, 视网膜血管分割

A Surrogate-Assisted U-Net Architecture Evolutionary Approach Based on GNN

Runjia Wu

School of Mathematics and Statistics, Guangdong University of Technology, Guangzhou Guangdong

Received: Feb. 28th, 2025; accepted: Mar. 27th, 2025; published: Apr. 3rd, 2025

Abstract

Retinal vascular sections play an important role in diagnosing and treatment of disease conditions. The technology based on recognition neural networks has made many breakthroughs in this field, but designing the optimal network structure is a challenging problem. Neural structure search faces challenges in real-world programs, such as too much search space and too long learning time. To effectively address these issues, this data demonstrates a U-Net architecture search method based on advanced algorithms. Unlike adding network parameters, this method can improve network

performance. Meanwhile, this architecture combines surrogate models and evolutionary algorithms to reduce computational costs during the search process. This article uses neural networks as an alternative model to enhance mutation algorithms and performs matching measures by drawing common objects. Meanwhile, this article accelerates the convergence speed of genetic computing methods. The experimental results show that the algorithm still outperforms state-of-the-art methods in key indicators such as sensitivity and AUROC.

Keywords

Surrogate Model, Evolutionary Algorithm, Neural Architecture Search, Retinal Vessel Segmentation

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

视网膜血管分割(Retinal Vessel Segmentation, RVS)是医疗影像分析中的重要分支,在临床诊断中具有重要应用价值。这项技术通过计算机辅助手段实现眼底血管网络的精确提取与量化分析,为眼科疾病的预防和治疗提供了可靠的客观依据。RVS在青光眼、高血压视网膜病变等多种眼底疾病的早期检测中具有显著优势。医学专家通常通过观察眼底微血管的形态特征来判断病情发展阶段及预后情况。鉴于传统人工分析方法效率低下且存在主观差异,基于人工智能的RVS算法开发已成为数字医疗领域的重要研究方向[1]。

在视网膜血管分析的初期阶段,主要采用基于边界识别的技术手段,其核心目标是区分眼底图像中的血管结构与周围组织。传统方法主要包含基于灰度值的分割算法和轮廓匹配策略。尽管梯度检测器如Sobel和Canny在血管识别中表现出一定效果,但往往难以区分真实血管与病灶区域或图像伪影,导致识别准确度受限[2]。轮廓匹配方法通过优化目标函数来拟合血管边缘,例如基于主动轮廓的检测模型,虽然具有较好的抗噪性能,但需要预先设定边缘特征参数和图像处理函数。

近年来,深度学习技术在医学影像分析领域取得突破性进展,逐渐成为血管分割的核心方法。这类方法通过计算每个像素点属于血管的概率来实现精确分割。其中,U型网络及其改进版本凭借优异的特征学习能力成为该领域的标准框架。该网络采用特征压缩与还原的双路径结构,有效融合了图像的全局语义和局部细节信息。研究者们通过多路径特征融合机制提升了模型性能,同时采用跨层连接策略增强了多尺度特征的表达能力[3]。然而,这些模型仍存在明显不足:网络结构依赖人工设计,参数量庞大,且容易过度拟合训练数据,导致泛化性能受限[4]。因此,如何在保证精度的前提下构建轻量化网络成为研究重点,循环卷积结构的引入为解决这一问题提供了新思路。

为减少人工设计网络的工作量,自动化网络结构搜索技术逐渐兴起。基于生物进化原理的优化算法在这一领域得到广泛应用。有研究者率先将该类算法应用于血管分割任务,成功发现了更精简的网络结构;另有学者通过改进的进化策略探索最优网络配置[5]。虽然这类方法展现出巨大潜力,但计算资源消耗大、搜索效率低等问题仍然突出。为此,研究者提出了基于性能预估的加速策略,通过构建替代模型来评估网络性能[6]。其中,图神经网络因其能够直观地表示网络拓扑结构,被广泛用于构建性能预测器,通常采用回归损失进行训练[7]。有研究表明,相比精确预测网络性能,准确评估不同结构的相对优劣更为重要。然而,现有方法多局限于小规模搜索空间,其在实际应用中的有效性仍需进一步验证[8]。

本文提出一种基于 GNN 代理模型的改进 U-Net 网络架构搜索方法。具体而言,我们以 U-Net 架构为基础,运用进化算法搜索块结构和操作。在前人研究基础上,循环卷积操作最大限度减少冗余参数。为快速找到适用于不同 RVS 任务的最优网络,我们借助基于 GNN 的代理模型加速进化算法搜索。大量实验表明,算法能在不同数据集中找到参数最少的最优网络。本研究主要贡献如下:

- 1) 采用基于 GNN 的代理模型,建立网络拓扑与特征映射关系,通过比较不同网络排名加速进化算法搜索。
- 2) 通过减少块中节点数量,找到参数最少且与最优算法性能相当的网络。引入递归卷积操作后,F1 分数准确性与其他模型相当。
- 3) 设计改进的自动架构搜索算法,在 DRIVE 和 CHASE DB1 两个数据集上与多种先进算法对比,结果表明该算法能快速准确地找到适用于不同 RVS 任务的最优网络,相关代码已上传至 GitHub。

2. 研究方法

在本节中,我们将详细阐述所提出的算法。首先介绍进化算法(EA)的编码方式及各种操作的设计,然后简要介绍图注意力网络(Graph Attention Network, GAT)。接着,我们提取网络架构,并利用 GAT 作为代理模型来预测网络排名的设计。

2.1. EA 设计

我们采用进化算法进行视网膜血管分割(RVS)架构搜索,维护一个包含 N 个个体的种群 $C = \{C_1, C_2, \dots, C_n\}$ 。在整个搜索空间 N 中,我们用一组二进制数 $C_i(a)$ 表示采样个体,目标是确保个体 $\max\{C\}$ 所代表的网络 N_i 为最优网络。

1) 搜索空间和编码

U-Net 网络通常由两部分构成:编码器 D 和解码器 E,二者均由若干个块组成。需注意,编码器的块数通常比解码器多一个。U-Net 通过跳跃连接(中间块除外)将这些块相连,使解码器能够利用编码器中提取的丰富语义信息重建输出图像。在算法中,解码器采用逐元素相加的方式整合跳跃连接和上采样特征。这种逐元素相加不仅效果良好,还减少了计算参数的数量。

每个块包含若干节点,节点之间的连接由进化算法的编码确定。其中,第一个节点的输入固定为前一个块的输出,最后一个节点的输出则是下一个块的输入。若中间节点无前驱节点,则将其连接到第一个节点;若无后继节点,则连接到最后一个节点;若既无前驱节点也无后继节点,则将该节点丢弃。

操作由基本操作单元构成,包括 3×3 卷积/ 5×5 卷积、实例归一化(Instance Normalization, IN)、修正线性单元/米什激活函数(ReLU/Mish)以及循环卷积。ReLU 是深度神经网络中广泛使用的激活函数之一,因其简单易用且计算成本低而备受青睐。Mish 是一种高性能的非单调、平滑且稳健的激活函数,在图像分类、目标检测和自然语言处理等任务中表现优于 ReLU 激活函数。实例归一化将图像每个实例的均值和方差归一化到同一水平,有助于模型分割图像中的不同对象。循环卷积在每个尺度上重复卷积,积累特征信息,使特征表示优于标准卷积。不同时间步的卷积将获得更具表现力的特征,有助于提取低级特征。循环卷积在不增加参数的情况下提高了网络的分割性能。

我们利用一组二进制数(即 0 和 1)来表示采样网络,将其划分为若干段,每段代表一个块,如图 1 所示。相应节点的操作由 4 位表示,0 和 1 代表选择的基本操作类型。所有基本操作按照位的顺序组合成一个操作(3×3 卷积/ 5×5 卷积 \rightarrow IN \rightarrow ReLU/Mish)。最后一位表示是否采用循环卷积方法,总共有 16 种不同的组合。每个段的最后几位表示节点的连接方式,0 和 1 分别表示是否连接到前驱节点。

2) 交叉和变异算子

在进化算法中,交叉操作通过重组不同个体的进化信息来生成具有不同特征的新个体。这些新个体

可能对环境变化具有更强的适应性，从而提高种群的稳健性。我们将交叉操作分为不同的粒度：单点交叉、多段交叉和多点交叉。显然，单点交叉生成的个体与父代差异最小，不适合搜索更优秀的个体。多点交叉随机选择多个点进行交叉操作，生成的个体与父代差异最大，可以提高种群的多样性。多段交叉则介于两者之间。根据 U-Net 网络的结构特征，仅对相应的块进行单点交叉操作，以确保其他未交叉块的完整性。种群中的个体以概率 P_{cross} 进行交叉操作。

在算法的早期阶段，我们采用多点交叉，一方面确保种群的多样性，另一方面为代理模型提供更全面的网络架构信息。在算法的后期阶段，采用多段交叉来加速算法的收敛。变异可以防止算法陷入局部最优。个体以 P_{mut} 的概率进行变异操作。在变异个体中，每个位以较低的概率 P_{flip} 翻转。

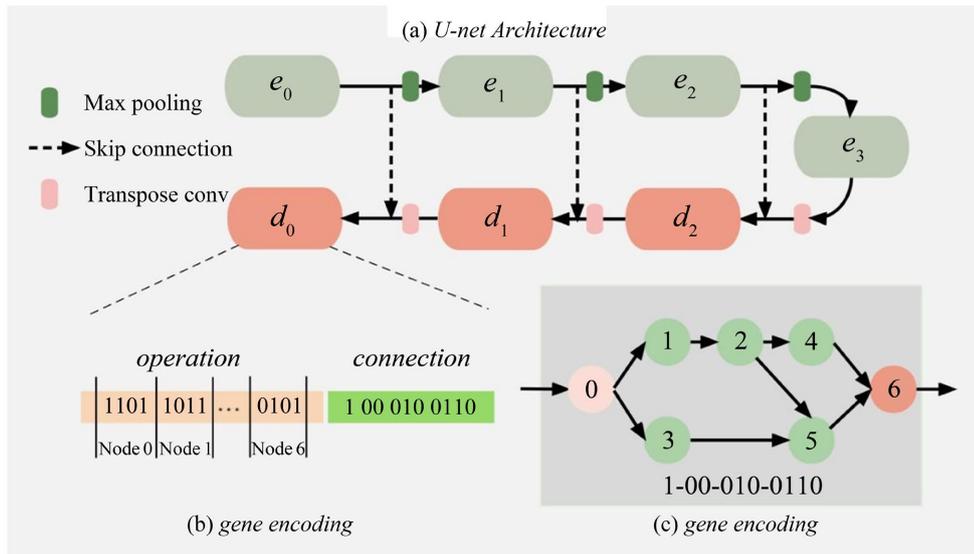


Figure 1. Algorithm architecture diagram
图 1. 算法架构图

2.2. 代理模型

代理模型旨在无需从头开始训练网络，即可对其性能进行估计。在进化算法中，代理模型可用于引导搜索方向，从而帮助进化算法更迅速地收敛至最优解。代理模型包含特征提取与性能预测两个部分。

1) 图注意力网络

图神经网络(Graph Neural Network, GNN)能够通过提取相应的特征向量进行学习和预测，从而捕捉图数据中节点和边之间的关系。其核心思想是将图数据表示为图结构，然后通过消息传递在图中传播信息。在每次消息传递步骤中，每个节点会聚合来自其相邻节点的信息并更新自身状态。这一过程会不断重复，直到节点的状态稳定，或者达到预设的迭代步数。基于图注意力机制的图神经网络(Graph Attention Network, GAT)是其中一种常见的类型。GAT 允许每个节点为其相邻节点的信息分配不同的权重，从而能够更有效地捕捉图中的长程依赖关系。

对于节点 i ，我们考虑其相邻节点 j ($j \in \bar{N}_i$, \bar{N}_i 表示节点 i 的一阶相邻节点集)的特征向量 h_j 。我们使用注意力权重 a_{ij} 来衡量和之间的关联程度。注意力权重通过以下公式计算：

$$a_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(a^T \left[Wh_i \parallel Wh_j\right]\right)\right)}{\sum_{k \in \bar{N}_i} \exp\left(\text{LeakyReLU}\left(a^T \left[Wh_i \parallel Wh_k\right]\right)\right)}$$

其中 a^T 、 W 是可训练参数。*LeakyReLU* 是带泄漏的修正线性单元激活函数。使用 *softmax* 函数对其相邻节点的注意力权重进行归一化,使得所有相邻节点的注意力权重之和为 1,表示节点 j 对节点 i 的相对重要性。然后,利用计算得到的注意力系数聚合特征,得到节点 i 新的特征表示:

$$h'_i = \sigma \left(\sum_{k \in \mathcal{N}_i} a_{ij} W h_j \right)$$

其中, σ 表示激活函数。通过重复上述聚合过程,得到最终的节点特征 h' 。随后我们采用平均聚合方法将所有节点的特征表示整合起来,作为整个图的特征表示。

在传统的 GNN 中,节点特征通常是利用节点标签进行训练的。然而,在代理模型 GNN 中,我们训练的是整个图的特征。为了更充分地利用节点特征信息,我们采用了嵌入方法来提取节点特征。此外,考虑到网络是以有向无环图的形式呈现的,节点的顺序会对网络的特征表示产生影响。因此,我们引入了类似于 Transformer 模型中的位置信息:

$$h_i = Emb_i + Pos_i$$

我们采用绝对位置编码的方式,为每个节点分配一个唯一的位置编码向量。该向量由正弦和余弦函数的线性组合构成,从而为每个节点提供了一个独特的空间位置标识。

2) 代理模型的性能预测

在训练性能预测器时,常用的损失函数(例如均方误差或 L1 损失)主要目的是最小化预测值与实际值之间的差距。然而,在网络排名的场景下,仅仅减少预测值与实际值的差异并不一定能提升预测器的性能。在网络排名中,网络的相对排名顺序决定了搜索算法选择保留哪些网络。这意味着即使预测器的准确率很高,也可能对排名顺序做出错误的预测,导致更优的网络个体被误淘汰。因此,我们提出了一种新的损失函数来替代传统的均方误差损失,使预测器能够预测网络的排名而非实际值。

3. 研究实验与分析

在本节中,我们首先对基于图注意力网络(GAT)的代理模型的有效性进行验证。之后,我们运用基于代理模型的进化算法(EA),在医学图像上搜索适用于视网膜血管分割任务的最优网络架构,并将其性能与其他最先进的算法进行对比分析。

为确保算法之间的可比性,我们使模型的基本参数与文献 1 中保持一致。代理模型的启用时间设定为 20,块节点的数量设定为 5。架构搜索将在 NVIDIA GeForce RTX3090 上进行。

3.1. 数据集和评估

1) 数据集

NAS-Bench-101: NAS-Bench-101 是神经架构搜索领域的一个基准数据集。它由多个单元组成的搜索空间构成,每个单元是一个由所有可能的具有 7 个节点(其中 2 个为输入/输出节点)的有向无环图(DAG)构成的网络。每个节点可以执行 3 种可能的操作(3×3 卷积、 1×1 卷积、 3×3 最大池化)。研究人员无需自行训练网络,可以直接调用该数据集中相应网络的预训练结果。该数据集将被用于验证代理模型的性能。

RVS 数据集: 我们将在两个公共数据集 DRIVE 和 CHASE DB1 对我们的算法进行验证。在实验中,我们选取这两个数据集中第一位专家的标注作为真实值的参考标准。DRIVE 数据集具备官方既定的训练与测试划分方式,其中 20 张图像被分配用于训练,另外 20 张图像则用于测试。而 CHASE DB1 数据集的划分比例为 20:8,即 20 张图像用于训练,剩余 8 张图像用于测试。我们计划在 DRIVE 数据集上

开展架构搜索，以确定最优网络架构，并进一步在 CHASE DBI 数据集上对所确定的最优网络架构的性能进行验证评估。

2) 评估指标

与采用均方误差作为损失函数的传统代理模型不同，我们的代理模型专注于评估网络的排名。鉴于此，我们选用肯德尔等级相关系数(KTau)作为评估指标。KTau 是一种广泛应用于衡量排名相关性的统计量，它能够有效量化两个排名之间的相似程度。对于我们的代理模型来说，KTau 相较于均方误差更具适用性，因为它能够精准地反映网络在排名方面的表现。

$$KTau = 2 * \frac{\text{一致对的数量} - 1}{n(n-1)/2}$$

在肯德尔等级相关系数(KTau)的计算中， n 表示样本数量。一致对是指在两个排名中，两个变量之间的顺序关系保持一致的情况。KTau 的取值范围为 $[-1, 1]$ ，用于评估预测排名的质量。该值越高，表明模型预测排名的准确性越高。

在视网膜血管分割(RVS)任务中，每个图像点被处理为一个二分类问题。因此，每个像素存在四种可能的结果：真正例(True Positive, TP)、假正例(False Positive, FP)、真反例(True Negative, TN)和假反例(False Negative, FN)。基于这些结果，RVS 任务为图像生成以下五个评估指标：

准确率(Accuracy, ACC)：衡量模型预测正确的比例。

灵敏度(Sensitivity, SE)：衡量模型正确识别正例的能力。

特异性(Specificity, SP)：衡量模型正确识别反例的能力。

F1 分数(F1 Score)：综合考虑精确率和召回率的指标。

受试者工作特征曲线下面积(Area Under the Receiver Operating Characteristic Curve, AUROC)：衡量模型在不同阈值下的综合性能。

3.2. 代理模型性能

为了验证代理模型的有效性，我们在不同比例的 NAS-Bench-101 数据集上对代理模型的参数进行了训练。我们随机抽取了 200 个不同的网络进行评估。可以明显看出，随着训练集比例的增加，我们的模型在测试集上的 K-Tau 指标呈现出不断上升的趋势。而模型在训练集上的 K-Tau 指标值出现下降，是因为当数据量过大时，模型难以从中学习到有用的知识。此外，我们的模型在测试集和训练集上的表现均优于 Peephole 和 E2EPP 模型。尽管在 0.1% 的数据集上，我们的模型表现不如 ReNAS，但作为代理模型，我们更关注其在少量训练数据下所展现出的高性能优势。显然，在实际应用中，我们不能等到进化算法完成之后才开始使用代理模型。

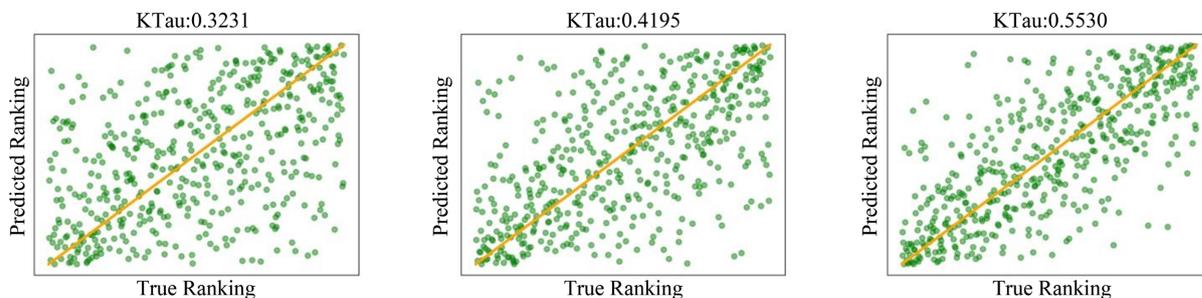


Figure 2. KTau results at different ratios

图 2. 不同比例下的 KTau 结果

与 NAS-Bench-101 不同, 我们设计的 U-Net 架构拥有一个更为复杂的搜索空间。我们从这个搜索空间中随机抽取了 1000 个网络。然后, 我们分别使用 200 个、300 个和 500 个不同数量的数据集来训练代理模型, 并利用剩下的 500 个网络进行测试, 测试结果如图 2 所示。在图 2 中, 每个点的 x 轴表示所有点的真实排名, 而 y 轴则对应于预测排名。理想情况下, 完美的预测结果应该位于直线 $y=x$ 上, 因此, 越接近直线 $y=x$ 的点就表示预测结果越好。从图中可以明显看出, 即使使用 500 个网络进行训练, 代理模型的 K τ 值也仅为 0.5530。这表明代理模型并非完美无误, 存在一定的误差, 这也凸显了设计评估指标的必要性。

3.3. 与最先进的 NAS 方法的比较

我们在两个公共数据集(DRIVE 和 CHASE DB1)上, 对所提出的算法进行了测试, 并将其与其他现有方法进行了比较。实验中, 我们使用了在 DRIVE 数据集上搜索到的最优网络。在不使用代理模型且其他参数保持一致的情况下, 整个搜索过程总共需要 5.5 个 GPU 天。而采用我们提出的方法, 仅需 3.7 个 GPU 天。这表明代理模型显著加快了算法的搜索速度。另一方面, 在 DRIVE 数据集上, 使用代理模型找到的最优网络在参数数量和 F1 分数方面均优于不使用代理模型的情况。综合来看, 代理模型在加速收敛和指导进化算法搜索方面发挥了重要作用。此外, 我们还在另一个数据集上对找到的最优模型的性能进行了验证。结果表明, 即使参数数量更少, 我们的网络也能与其他最先进的算法相媲美, 并且在不同数据集上均取得了最佳性能。

4. 结论

在本研究中, 我们提出了一种改进的新型神经架构搜索方法, 用于视网膜血管分割任务。算法借助基于图神经网络(GNN)的代理模型来辅助进化算法搜索最优网络架构。我们对基于架构的参数数量进行了限制, 使其保持在一个较小的范围内。通过引入循环卷积操作, 在不增加额外参数的情况下, 有效提升了网络的分割性能。实验结果表明, 算法所找到的最优架构在某些关键指标上优于当前主流的视网膜分割方法。这一发现为未来视网膜分割领域中小模型的研究提供了宝贵的参考和领域知识。

参考文献

- [1] Bulut, M., Kurtuluş, F., Gözkaya, O., Erol, M.K., Cengiz, A., Akıdan, M., et al. (2017) Evaluation of Optical Coherence Tomography Angiographic Findings in Alzheimer's Type Dementia. *British Journal of Ophthalmology*, **102**, 233-237. <https://doi.org/10.1136/bjophthalmol-2017-310476>
- [2] Zhu, L., Zong, Y., Yu, J., Jiang, C., He, Y., Jia, Y., et al. (2018) Reduced Retinal Vessel Density in Primary Angle Closure Glaucoma: A Quantitative Study Using Optical Coherence Tomography Angiography. *Journal of Glaucoma*, **27**, 322. <https://doi.org/10.1097/jgg.0000000000000900>
- [3] Mahiba, C. and Jayachandran, A. (2019) Severity Analysis of Diabetic Retinopathy in Retinal Images Using Hybrid Structure Descriptor and Modified CNNs. *Measurement*, **135**, 762-767. <https://doi.org/10.1016/j.measurement.2018.12.032>
- [4] Radha, K. and Karuna, Y. (2023) Retinal Vessel Segmentation to Diagnose Diabetic Retinopathy Using Fundus Images: A Survey. *International Journal of Imaging Systems and Technology*, **34**, e22945. <https://doi.org/10.1002/ima.22945>
- [5] Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M. and Goldbaum, M. (1989) Detection of Blood Vessels in Retinal Images Using Two-Dimensional Matched Filters. *IEEE Transactions on Medical Imaging*, **8**, 263-269. <https://doi.org/10.1109/42.34715>
- [6] Abdou, I.E. and Pratt, W.K. (1979) Quantitative Design and Evaluation of Enhancement/thresholding Edge Detectors. *Proceedings of the IEEE*, **67**, 753-763. <https://doi.org/10.1109/proc.1979.11325>
- [7] Ilayarajaa, K.T. and Logashanmugam, E. (2020). Retinal Blood Vessel Segmentation Using Morphological and Canny Edge Detection Technique. 2020 *International Conference on System, Computation, Automation and Networking (ICSCAN)*, 3-4 July 2020, 1-5. <https://doi.org/10.1109/icscan49426.2020.9262446>
- [8] Ghadiri, F., Akbarzadeh-T, M. and Haddadan, S. (2011). Vessel Segmentation Based on Sobel Operator and Fuzzy Reasoning. 2011 *1st International eConference on Computer and Knowledge Engineering (ICCKE)*, 13-14 October 2011, 189-194. <https://doi.org/10.1109/iccke.2011.6413349>