

中文摘要

基于中性粒细胞形态构建识别骨髓增殖性肿瘤的深度学习模型

费城染色体阴性的骨髓增殖性肿瘤(myeloproliferative neoplasm, MPN) 是一组造血干细胞克隆性疾病,通常会导致一系或多系血细胞数量异常增高,主要包括原发性血小板增多症(Essential thrombocytosis, ET), 真性红细胞增多症(Polycythemia Vera, PV)和原发性骨髓纤维化(Primary Thrombocytosis, PMF) 。目前 MPN 主要依据临床表现、血液学改变、组织病理及出现典型的基因突变等来诊断。最近很多文献已经讨论了 MPN 的发生发展与炎症密切相关,甚至有学者认为 MPN 中炎症对造血干细胞的损伤早于基因突变的发生。Dragana 等人^[1]强调了中性粒细胞在 MPN 发生发展的重要性。此外,目前的研究提示肿瘤患者中存在三种异质性中性粒细胞。近几年来人工智能迅速发展,这一前沿技术也逐渐被应用于医学领域各个方面。人工智能中的深度学习拥有无与伦比的信息挖掘能力,无需依靠人为选择特征能够自动的去抓取潜在信息。在图像分类、目标检测及语义分割发挥了重要作用。越来越多的研究也证实了深度学习算法在血液病中鉴别不同细胞形态精确度和灵敏度可高达 90%以上,达到甚至超越专家水平。本研究提出了基于中性粒细胞形态构建识别骨髓增殖性肿瘤的深度学习模型判断患者是否为 MPN 以及 MPN 亚型新方法。我们选择的 PicoDet 深度学习目标检测方法达到了很好的分类预测效果。因此深度学习可以实现辅助 MPN 的诊断及分型。

关键词:

骨髓增殖性肿瘤, 中性粒细胞形态, 人工智能, 深度学习, 目标检测

Abstract

Construction of a deep learning model for recognizing myeloproliferative neoplasm based on neutrophil morphology

Ph- myeloproliferative neoplasm (MPN) is a group of hematopoietic stem cell clonal diseases. It usually causes abnormal increase of cell number of one or multiple lines. They mainly include Essential thrombocytosis (ET), Polycythemia Vera (PV) and Primary Thrombocytosis (PMF). At present, MPN is mainly diagnosed based on clinical manifestations, hematological changes, histopathology and typical gene mutations. Recently, many literatures have discussed that the occurrence and development of MPN is closely related to inflammation, and some scholars even believe that the damage of inflammation on hematopoietic stem cells in MPN is earlier than the occurrence of gene mutation. Dragana et al.^[1] emphasized the importance of neutrophils in the occurrence and development of MPNS. In addition, the present study suggests the presence of three heterogeneous neutrophils in tumor patients. With the rapid development of artificial intelligence in recent years, this cutting-edge technology has been gradually applied to various aspects of the medical field. Deep learning in artificial intelligence has unparalleled information mining ability. It can automatically capture potential information without relying on human selection features. It plays an important role in image classification, object detection and semantic segmentation. More and more studies have also confirmed that the accuracy and sensitivity of deep learning algorithm in distinguishing different cell morphology in blood diseases can reach more than 90%, reaching or even surpassing the expert level. In this study, a new method was proposed to establish a deep learning model based on neutrophil morphology to identify myeloproliferative neoplasm and determine whether patients are MPN and MPN subtypes. The PicoDet deep learning object detection method selected by us achieves good classification prediction effect. Therefore, deep learning can realize auxiliary diagnosis and classification of MPNS.

Key words

Myeloproliferative neoplasm, Neutrophil morphology, Artificial Intelligence ,
Deep learning, Object detection

目 录

第 1 章 引言	1
第 2 章 文献综述	3
2.1 骨髓增殖性肿瘤现状及研究进展	3
2.1.1 骨髓增殖性肿瘤现状	3
2.1.2 MPN 亚型可能与中性粒细胞相关	5
2.2 中性粒细胞在 MPN 中的作用	6
2.2.1 MPN 与炎症密切相关	6
2.2.2 中性粒细胞在 MPN 的发生发展起着重要作用	7
2.3 中性粒细胞形态存在异质性	8
2.3.1 中性粒细胞在实体瘤中有着不同的亚型	8
2.3.2 中性粒细胞形态可辅助血液病的诊断	9
2.4 深度学习在医学的应用	10
2.4.1 人工智能及深度学习定义	10
2.4.2 卷积神经网络 (CNN)	10
2.4.3 目标检测算法	12
2.4.4 深度学习在医学领域应用广泛	13
第 3 章 材料及方法	15
3.1 纳入标准	15
3.2 一般资料	15
3.3 实验设计	15
3.3.1 数据集描述	15

3.3.2 选择目标检测模型	18
3.3.3 模型验证与评价	18
第 4 章 结果	20
第 5 章 讨论	22
第 6 章 结论	25
参考文献	26
作者简介及在学期间所取得的成果	36
致 谢	37

英文缩略词

英文缩写	英文全称	中文名称
MPN	myeloproliferative neoplasm	骨髓增殖性肿瘤
PV	Polycythemia Vera	真性红细胞增多症
ET	Primary Thrombocytosis	原发性血小板增多症
PMF	Primary Thrombocytosis	原发性骨髓纤维化
TANS	Tumor associated neutrophils	肿瘤相关中性粒细胞
LDN	Low density neutrophils	低密度中性粒细胞
HDN	High density neutrophils	高密度中性粒细胞
MDSCs	Myeloid-derived suppressor cells	骨髓源性抑制细胞
HSCs	Hematopoietic stem cells	造血干细胞
HPCs	Hematopoietic progenitor cells	造血祖细胞
ROS	Reactive oxygen species	活性氧
RSN	reactive nitrogen species	活性氮
AI	Artificial Intelligence	人工智能
CNN	Convolutional Neural Networks	卷积神经网络
DL	Deep learning	深度学习
P	Precision	精度
R	Recall	召回率
AP	Average Precision	平均精度
mAP	mean Average Precision	各类别平均精度
TP	True positive	真阳性
TN	True negative	真阴性
FP	False positive	假阳性
FN	False negative	假阴性

第 1 章 引言

费城染色体阴性的骨髓增殖性肿瘤是一组造血干细胞恶性克隆性疾病，导致一系或多系的血细胞数量异常增高。MPN 主要包括原发性血小板增多症，真性红细胞增多症和原发性骨髓纤维化^[2]。

目前其发生主要与 JAK2、CALR、MPL 等驱动基因或者 TET2、DNMT3 非驱动基因突变相关。但是越来越多的研究者认为炎症及免疫功能失调在其发生发展过程中起着重要作用，甚至有学者强调了 MPN 中的中性粒细胞是独特的、特异性的炎症反应细胞，具有不可否认的恶性潜能^[1]。

MPN 的症状和临床表现重叠，疾病有时很难区分。且同样的突变可以导致 2 种或者 3 种不同的临床表型，即不能通过基因将其完全区分。体外克隆分析^[3]、动物模型^[4, 5]、人类 JAK2 V617F 等位基因负荷的定量研究^[6-8]以及其他突变基因^[9]患者中发现的特征表明基因负荷、骨髓微环境和非驱动基因突变都在一定程度上解释了 MPN 表型异质性。而骨髓微环境和基因突变都与中性粒细胞密切相关。

最近的研究提示肿瘤患者中的中性粒细胞存在异质性，至少存在肿瘤相关中性粒细胞(Tumor associated neutrophils, TANS)、髓源性抑制细胞(Myeloid-derived suppressor cells, MDSCs)和正常中性粒细胞三种不同功能的中性粒细胞。肿瘤相关中性粒细胞包括 N1 型中性粒细胞和 N2 型中性粒细胞两种不同功能和形态的群体，N1 型中性粒细胞抗肿瘤发生，N2 型中性粒细胞促进肿瘤的发生。N1 型中性粒细胞形态较 N2 型中性粒细胞相比更具有超节段结构。髓源性抑制细胞是具有抑制 T 细胞激活和增殖能力的髓系细胞的异质亚群，其促进肿瘤的发生。有报道称 MDSCs 的异常增殖是骨髓增殖性肿瘤免疫失调的重要表现之一。Wang 等^[10]研究发现骨髓增殖性肿瘤患者的外周血 MDSCs 水平显著升高。目前还没有明显的标记物去鉴别不同的中性粒细胞，但是 Fridlender 等^[11]通过 MDSCs、TANS 及幼稚的中性粒细胞的转录组分析证实了它们是明显分开的群体。

目前人工智能发展迅速，并逐渐诞生了深度学习这一概念。深度学习计算量较大且拥有强大的信息挖掘能力，广泛应用于各个领域，其在血液病的应用也经受了考验。Jan-Niklas 等人基于深度学习发现了有 NPM1 突变的急性髓系白血

病中的中性粒细胞目前尚未报道的细胞形态学特征^[12]。我们前期的研究也发现了中性粒细胞在不同血液病中有着不同的形态^[13, 14]。鉴于中性粒细胞在 MPN 发生发展中扮演了重要角色及中性粒细胞存在异质性,同时得益于深度学习在血液病领域具有卓越的性能。我们推测基于中性粒细胞形态学建立的深度学习模型可识别 MPN 的不同亚型,达到辅助 MPN 诊断及鉴别诊断的目的。

本研究建立了一个 4800 张 MPN 中性粒细胞数据集,采用 PicoDet 目标检测模型去鉴别中性粒细胞是否为 MNP 中性粒细胞以及是 PV 中性粒细胞还是 ET 中性粒细胞,从而辅助 MPN 诊断及鉴别诊断。这将极大的节约诊疗时间、经济,造福广大患者。甚至使诊疗水平有限的医院对疑似 MPN 患者进行初步筛查,以便患者是否需进一步至上级医院诊治。

第2章 文献综述

2.1 骨髓增殖性肿瘤现状及研究进展

2.1.1 骨髓增殖性肿瘤现状

费城染色体阴性的骨髓增殖性肿瘤是一种恶性克隆性血液学疾病,经常表现为外周血血细胞一系或多系增多,常伴肝脾肿大,最后进展为骨髓纤维化、骨髓衰竭甚至向急性白血病转化^[15, 16]。其主要包括原发性血小板增多症,真性红细胞增多症和原发性骨髓纤维化。而原发性骨髓纤维化又分为了纤维化前期 PMF 和明显纤维化期 PMF。MPN 的发生主要与 JAK2V617F、CALR 及 MPL 驱动基因突变激活 JAK-SATA 通路相关。但是随着不断对 MPN 的研究,越来越多的学者认为非驱动基因突变、炎症及免疫功能失调在其发生发展过程中也起着重要作用。

骨髓增殖性肿瘤发病平均年龄在 60 岁左右。MPN 通常在男性中更常见,但在原发性血小板增多症中女性更为普遍^[17]。MPN 发病率较低,生存期较长^[18]。目前报道的 MPN 发病率从 1.15/100000 到 4.99/100000 不等^[19], ET 患者生存期最长,可达 20 年, PV 患者生存期可达 14.1 年^[20]。纤维化前期 PMF 患者的中位生存期约为 14.7 年,但不良核型或突变型患者的生存期明显缩短^[21]。明显纤维化期 PMF 的预后较差,中位生存期为 3.1-5.8 年^[21]。MPN 各亚型的临床表现相似,主要有疲劳、皮肤瘙痒、盗汗、骨痛、发热和体重减轻等症状。

世界卫生组织(WHO)认为结合临床、形态学和分子遗传学特征是定义诸如 MPN 等疾病的最合适的方法,且强调了细胞形态在骨髓增殖性肿瘤的诊断重要性^[2, 22]。目前 MPN 的诊断主要依据临床表现、外周血细胞计数及血细胞比容、骨髓象改变、骨髓病理以及出现典型的基因突变等。明确是否有驱动基因突变在诊断 MPN 时尤为重要,而骨髓活检在鉴别 MPN 亚型中有着不可替代的作用。2016 年 WHO 提出了新的诊断依据具体见下表 2.1。

表 2.1 MPN 2016 年 WHO 诊断标准^[2]

Polycythemia vera (PV) ^a		Essential thrombocythemia (ET) ^b	
Major criteria			
1 Hemoglobin > 16.5 g/dL (men) Hemoglobin > 16.0 g/dL (women) or Hematocrit > 49% (men) Hematocrit > 48% (women) or increased red cell mass (RCM) ^c		Platelet count ≥ 450 × 10 ⁹ /L	
2 BM biopsy showing hypercellularity for age with trilineage growth (panmyelosis) including prominent erythroid, granulocytic and megakaryocytic proliferation with pleomorphic, mature megakaryocytes (differences in size)		BM biopsy showing proliferation mainly of the megakaryocyte lineage with increased numbers of enlarged, mature megakaryocytes with hyperlobulated nuclei. No significant left-shift of neutrophil granulopoiesis or erythropoiesis and very rarely minor (grade 1) increase in reticulin fibers ^d	
3 Presence of JAK2 or JAK2 exon 12 mutation		Not meeting WHO criteria for BCR-ABL1 + CML, PV, PMF, MDS, or other myeloid neoplasms	
4		Presence of JAK2, CALR or MPL mutation	
Minor criteria			
1 Subnormal serum erythropoietin level		Presence of a clonal marker (e.g., abnormal karyotype) or absence of evidence for reactive thrombocytosis	

Primary myelofibrosis (PMF) ^a	
Prefibrotic/early PMF (pre-PMF)	Overt PMF
Major criteria	
1 Megakaryocytic proliferation and atypia ^b , without reticulin fibrosis > grade 1 ^c , accompanied by increased age-adjusted BM cellularity, granulocytic proliferation and often decreased erythropoiesis	Megakaryocyte proliferation and atypia ^b accompanied by either reticulin and/or collagen fibrosis (grade 2 or 3)
2 Not meeting WHO criteria for BCR-ABL1 + CML, PV, ET, MDS, or other myeloid neoplasm	Not meeting WHO criteria for BCR-ABL1 + CML, PV, ET, MDS or other myeloid neoplasm
3 Presence of JAK2, CALR or MPL mutation or in the absence of these mutations, presence of another clonal marker ^d or absence of minor reactive BM reticulin fibrosis ^e	Presence of JAK2, CALR or MPL mutation or in the absence, the presence of another clonal marker ^d or absence of evidence for reactive BM fibrosis ^f
Minor criteria	
1 Presence of one or more of the following, confirmed in two consecutive determinations: • Anemia not attributed to a comorbid condition • Leukocytosis ≥ 11 × 10⁹/L • Palpable splenomegaly • LDH level above the upper limit of the institutional reference range	Presence of one or more of the following confirmed in two consecutive determinations: • Anemia not attributed to a comorbid condition • Leukocytosis ≥ 11 × 10⁹/L • Palpable splenomegaly • LDH level above the upper limit of the institutional reference range • Leukoerythroblastosis

不同亚型的骨髓增殖性肿瘤有着各自的特点。PV 的特征是红细胞的产生增加明显，包括血红蛋白浓度和红细胞压积增加。其往往会伴随着血小板和/或白细胞增多。绝大多数携带 JAK2V617F 突变，占全部病例的 95%以上，少数具有 JAK2 第 12 外显子突变。ET 的特征是持续的血小板增多和巨核细胞数量的增多，一般不伴有血小板和白细胞的增多。其最常见的驱动突变是 JAK2V617F，约占全部病例的 50-60%，CALR 突变约占全部病例的 30%，MPL 突变约占全部病例的 5%。PMF 的特征是异常的巨核细胞和粒细胞的增殖，以及骨髓中纤维化组织的沉积和髓外造血功能。其中 JAK2V617F 突变占全部病例的 50-60%，CALR 突

变占全部病例的 30%，MPL 突变占全部病例的 8%。骨髓病理在 MPN 的鉴别诊断起着重要的作用，尤其是在原发性血小板增多症和纤维化前期 PMF 中的鉴别。原发性血小板增多症的骨髓增生程度正常，以巨核细胞增生为主，粒系及红系增生正常。通常也没有核左移。巨核细胞一般随机散在分布，细胞核的体积较大，可呈过分叶。而纤维化前期 PMF 患者骨髓极度增生，以粒系及巨核系增生为主，红系一般轻至中度减低。巨核细胞一般集中在一起，细胞大小不一，细胞核分叶少。真性红细胞增多症的患者骨髓病理可以三系增生，红系、粒系及巨核系都可增生。巨核细胞形态通常为多形性成熟的细胞。

骨髓涂片及活检病理检查取决于病理科医生的知识、经验且比较费时。另外大部分医院暂未开展 NGS 基因突变检测这一工作，医务工作者往往将标本送往第三方检测公司，进一步耗费患者诊治时间。

2.1.2 MPN 亚型可能与中性粒细胞相关

虽然 JAK2V617F、CALR 及 MPL 驱动基因突变是骨髓增殖性肿瘤发生的主流观点，但是同样的突变可以导致三种不同的临床亚型，甚至无驱动基因的突变也可以导致 MPN 的发生，也就说无法单独从基因方面进行 MPN 分型。

针对同样的突变导致不同的临床表型这个问题，来自全球的学者通过不断的努力发现了等位基因负荷、突变型 JAK2V617F 与野生型 JAK2 的比率、非驱动基因突变以及体细胞突变顺序、CALR 突变的类型及骨髓微环境都在一定程度上解释了此问题。

研究发现，等位基因负荷与 MPN 疾病表型有关^[8]，JAK2V617F 突变存在纯合和杂合 2 种类型。ET 几乎都是杂合子，而 PV 72%为杂合子，28%为纯合子，PMF 86%是杂合子，14%为纯合子。

Ralph Tiedt 等人^[23]发现了真性红细胞增多症中突变型 JAK2V617F 与野生型 JAK2 的比率明显高于原发性血小板增多症，提示了突变型 JAK2V617F 与野生型 JAK2 的比率在一定程度上决定了转基因小鼠的 MPN 表型。

据报道，TET2 和 DNMT3A 是 MPN 患者除驱动基因突变以外常见的基因突变且可能跟 MPN 亚型相关。TET2 功能的丧失会导致 DNA 超甲基化，使造血细胞偏向骨髓单核细胞分化^[24]。一项研究报道了 JAK2V617F 阳性的 MPN 模型中

DNMT3A 的功能丧失会阻止红系的增殖。也有研究发现体细胞突变顺序可能会导致不同的临床亚型, ORTMANN 等^[25]发现了 TET2 突变早于 JAK2V617F 发生的患者倾向于 ET 表型, 而 TET2 突变在 JAK2V617F 之后的患者倾向于 PV 表型。

CALR 突变的类型也提供了一些 MPN 表型差异, CALR1 型突变在原发性骨髓纤维化中更常见, 相比之下, CALR2 型突变更容易导致 ET, 并且比 CALR 1 型突变的 ET 有更高的血小板计数^[23]。也有研究表明骨髓微环境也可能是 MPN 表型异质性的影响因素^[26]。而这些基因突变、骨髓微环境与长期炎症的刺激和中性粒细胞损伤造血干细胞密切相关。

2.2 中性粒细胞在 MPN 中的作用

2.2.1 MPN 与炎症密切相关

近年来, 炎症促进 MPN 的发生被越来越多的学者接受。Hasselbalch 等人收集了流行病学、生物学、病理学和临床数据, 认为 MPN 是炎症性疾病, 而且可以被看做是慢性炎症和肿瘤发生之间联系的范例^[27], 并提议了 MPN 可以被认为是一种“人类炎症模型”^[28]。Laura F 等^[29]阐明了炎症在 MPN 中创造了一个有利于 JAK2V617F 肿瘤性造血干细胞生长的环境。事实上, 慢性炎症是肿瘤的一个重要的启动事件^[30, 31], 它可通过氧化应激机制驱动克隆进化^[32, 33]。MPN 中慢性炎症和克隆进化之间的联系已经被 TNF- α 促进 JAK2V617f 阳性细胞克隆扩增所证实^[34]。

MPN 病理生理与氧化应激之间的关系已经在很多研究中被反复强调了^[35-38]。研究表明活性氧 (Reactive oxygen species, ROS) 产生的增加和活性氮 (reactive nitrogen species, RSN) 相关参数的升高与 MPN 发病密切相关^[38]。ROS 增强可能抑制磷酸酶活性, 导致磷酸化依赖性的 JAK/STAT 信号延长^[39]。ROS 和 RNS 水平的增加可以直接导致干细胞内的 DNA 损伤^[40], ROS 的升高已被证明会降低星状细胞的自我更新能力, 使星状细胞脱离静止状态^[41, 42]。在炎症介导的髓系增生过程中, 携带 DNA 损伤的髓系细胞可能最终获得突变和基因组不稳定, 这既可能启动克隆性骨髓增生, 也可能推动克隆进化走向骨髓纤维化阶段^[43, 44]。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/136153054231010055>