



乳腺钼靶肿块自监督预训练迁移检测方法研究

莫宏伟, 孙琪, 孙鹏, 张显玉, 孙江宏, 孙惟嘉

引用本文:

莫宏伟, 孙琪, 孙鹏, 张显玉, 孙江宏, 孙惟嘉. 乳腺钼靶肿块自监督预训练迁移检测方法研究[J]. 智能系统学报, 2024, 19(5): 1082-1091.

MO Hongwei, SUN Qi, SUN Peng, et al. Self-supervised pretraining detection of mammographic mass targets in breast[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(5): 1082-1091.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202304032>

您可能感兴趣的其他文章

样本仿真结合迁移学习的声呐图像水雷检测

Detection of underwater mine target in sidescan sonar image based on sample simulation and transfer learning
智能系统学报. 2021, 16(2): 385-392 <https://dx.doi.org/10.11992/tis.202101030>

基于二进制生成对抗网络的视觉回环检测研究

Visual loop closure detection based on binary generative adversarial network
智能系统学报. 2021, 16(4): 673-682 <https://dx.doi.org/10.11992/tis.202007007>

一种基于2D时空信息提取的行为识别算法

A behavioral recognition algorithm based on 2D spatiotemporal information extraction
智能系统学报. 2020, 15(5): 900-909 <https://dx.doi.org/10.11992/tis.201906054>

半监督自训练的方面提取

Aspects extraction based on semi-supervised self-training
智能系统学报. 2019, 14(4): 635-641 <https://dx.doi.org/10.11992/tis.201806006>

基于深度学习的椎间孔狭窄自动多分级研究

Deep learning based automatic multi-classification algorithm for intervertebral foraminal stenosis
智能系统学报. 2019, 14(4): 708-715 <https://dx.doi.org/10.11992/tis.201806015>

基于自编码器的特征迁移算法

Feature transfer algorithm based on an auto-encoder
智能系统学报. 2017, 12(6): 894-898 <https://dx.doi.org/10.11992/tis.201706037>

DOI: 10.11992/tis.202304032

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20240830.1535.002>

乳腺钼靶肿块自监督预训练迁移检测方法研究

莫宏伟¹, 孙琪¹, 孙鹏¹, 张显玉², 孙江宏³, 孙惟嘉³

(1. 哈尔滨工程大学智能科学与工程学院, 黑龙江 哈尔滨 150001; 2. 哈尔滨医科大学附属肿瘤医院 乳腺外科, 黑龙江 哈尔滨 150001; 3. 哈尔滨医科大学附属肿瘤医院 影像中心, 黑龙江 哈尔滨 150001)

摘要: 借助深度学习技术在乳腺钼靶领域辅助医生进行乳腺癌诊断在当下已经成为很多研究关注的热点, 诊断技术主要包括良恶性分类、病灶区域检测以及病灶区域分割等。由于深度学习训练的模型性能很大程度上依赖于大量的带有标注的数据, 而医学图像数据集往往存在数据量少、标注成本昂贵以及公开数据集标注质量差等现象, 所以在医学图像领域应用深度学习技术具有重重困难。为使基于深度学习的乳腺钼靶计算机辅助诊断技术的开发不受限于大量有标注的数据, 提出一种适用于钼靶自监督目标检测方法来完成乳腺钼靶肿块检测任务, 利用大量来自肿瘤医院的数据预训练, 并在公开数据集 DDSM 上进行微调与测试。实验结果表明, 提出模型在乳腺钼靶肿块检测任务中表现优异, 并且不依赖于位置标签, 具有重要的研究价值与应用前景。

关键词: 目标检测; 自监督; 钼靶影像; 预训练; 数据增强; 视觉表示; 卷积神经网络; 图像分类

中图分类号: TP181 **文献标志码:** A **文章编号:** 1673-4785(2024)05-1082-10

中文引用格式: 莫宏伟, 孙琪, 孙鹏, 等. 乳腺钼靶肿块自监督预训练迁移检测方法研究 [J]. 智能系统学报, 2024, 19(5): 1082-1091.

英文引用格式: MO Hongwei, SUN Qi, SUN Peng, et al. Self-supervised pretraining detection of mammographic mass targets in breast[J]. CAAI transactions on intelligent systems, 2024, 19(5): 1082-1091.

Self-supervised pretraining detection of mammographic mass targets in breast

MO Hongwei¹, SUN Qi¹, SUN Peng¹, ZHANG Xianyu², SUN Jianghong³, SUN Weijia³

(1. College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China; 2. Breast Surgery Department, Cancer Hospital Affiliated to Harbin Medical University, Harbin 150001, China; 3. Imaging Center, Cancer Hospital Affiliated to Harbin Medical University, Harbin 150001, China)

Abstract: Deep learning technology has become a major research focus in assisting doctors with breast cancer diagnosis, particularly in mammographic imaging. This research encompasses tasks such as benign and malignant classification, lesion region detection, and lesion region segmentation. However, the performance of deep learning models heavily relies on large amounts of labeled data. Medical image datasets often suffer from limited data, high labeling costs, and poor labeling quality of open datasets, which make it challenging to apply deep learning technology effectively in the medical imaging field. To advance computer-aided diagnosis of breast mammographic targets using deep learning without relying on large amounts of labeled data, this study proposes a method for self-supervised target detection in mammographic images. The method addresses the task of detecting breast mammographic masses. Extensive data from cancer hospitals were pretrained, fine-tuned, and tested on the open dataset DDSM. The experimental results show that the proposed model performs excellently in detecting breast mammographic masses and does not rely on location tags. These results highlight the significant research value of the model and promising application prospects.

Keywords: object detection; self-supervised; mammographic; pretraining; data augment; visual representation; convolution neural network; image classification

随着深度学习技术的发展, 医学领域的计算机辅助诊断系统已经在放射学^[1-2]、皮肤科^[3]和病

理学^[4]等领域取得了令人瞩目的进展。乳腺癌的早期筛查主要通过 2 种成像方式: 乳腺的超声影像和钼靶影像。乳腺钼靶成像检查虽然具有一定的放射性, 但对乳腺内的微钙化比较敏感, 微钙

收稿日期: 2023-04-15. 网络出版日期: 2024-09-02.

通信作者: 莫宏伟. E-mail: honwei2004@126.com.

©《智能系统学报》编辑部版权所有

化是乳腺癌的特征之一, 所以钼靶成像检查结果可以作为乳腺癌早期筛查的黄金标准之一。但是由于乳腺钼靶图像分辨率高、病灶微小, 乳腺钼靶图像的阅片对医生的经验要求较高, 筛查结果通常也会受医生主观因素的影响。随着人工智能相关理论的发展, 基于深度学习的乳腺超声肿瘤识别技术有着重要的研究价值和意义^[5]。

目前的乳腺钼靶的异常检测方法分为传统方法和深度学习方法。传统方法通常包括 2 个步骤: 1) 手工设计钼靶图像中异常区域的特征; 2) 将特征传递给分类器, 通过训练分类器来检测异常区域。Dhungel 等^[6]提出了一种利用手工制作的特征对深度网络进行微调的乳腺肿块良恶性分类方法。深度网络通过预先训练手工制作的特征, 并在此基础上对深度网络进行微调。Kriti 等^[7]首先从肿块中提取纹理特征, 再通过主成分分析 (principal component analysis, PCA) 方法对肿块的特征空间进行降维, 最后选择几个主要特征进行分类。Ketabi 等^[8]提出了一种通过光谱聚类和支持向量机自动检测钼靶图像中的肿块区域的方法。使用光谱聚类, 通过基于区域直方图的强度描述符和基于灰度共生矩阵 (grey level co-occurrence matrix, GLCM) 的纹理描述符, 将感兴趣区域分割为标记的多截面模式, 并从分割的截面中导出形状和概率特征, 将其交给遗传算法 (genetic algorithm, GA) 进行特征选择。这些传统方法中, 特征在提取过程中需要依靠专家来设计, 即使是同一类型的病灶, 在形状、纹理等方面也有很大差异, 而且手工制作的特征并不能很好地描述不同病灶。

在深度学习技术飞速发展后, 出现了一些基于卷积神经网络的乳腺钼靶肿块检测方法。Sanasi 等^[9]也开发了用于将乳腺钼靶分类为良性或恶性的卷积神经网络, 将深度学习的概念与极限学习机 (extreme learning machine, ELM) 相结合, 该方法使用简单的群体搜索算法进行了优化。Gnanasekaran 等^[10]融合了公开数据集以及从医院得到的受限数据集, 同样使用卷积神经网络来实现乳腺钼靶图像良恶性的分类, 所提出的卷积神经网络模型由 8 个卷积层、4 个最大池化层和 2 个全连接层组成, 结果优于预先训练的 AlexNet 和 VGG16 网络。Al-masni 等^[11]开发了一种基于 YOLO 的计算机辅助诊断 (computer-aided diagnosis, CAD) 系统。所提出的系统可以在处理整个乳房图像的同时进行目标检测和图像分类。基于深度学习的方法不需要手工设计特征, 减轻了工作

量的同时又可以使网络学习到更通用的特征, 泛化性能更强。然而基于深度学习的检测方法需要大量的有标注数据集进行网络的训练, 而医学图像数据集往往规模较小, 并且标注昂贵且难以获得, 所以使用医学图像数据对网络从头开始训练不仅训练速度慢, 而且效果不理想。解决数据集不充分的常用手段之一是迁移学习^[12]。与医学影像分割任务不一样的是, 医学影像检测任务不需要对医学影像进行逐像素的分类, 只需要确定感兴趣对象的边界^[13]。

所以本文提出一种自监督预训练迁移检测模型, 所提出的模型能够较好地对图像的空间特征进行保留, 同时可以使用大量无标注的数据对特征提取网络进行训练, 再用少部分有标注的数据在检测任务中对检测模型中的检测组件进行训练和微调。实验表明, 本文提出的改进后的模型相比于其他目标检测模型而言, 对不同类别的目标的检测效果得到了全面的提升, 尤其对尺寸较小的肿块检测更加突出。不依赖于大量的有标注的数据集也可以达到甚至超越在大规模自然图像数据集上预训练的网络的效果。

1 本文提出的自监督检测框架

本节详细介绍了本文提出的自监督检测框架的设计思路与改进措施。

1.1 乳腺钼靶肿块自监督检测框架

本文提出一种自监督预训练迁移检测网络模型。模型整体的结构如图 1 所示, 模型分为 2 个部分: 预训练部分与检测部分。其中预训练部分的具体模型结构由 1.2 节给出, 检测部分的具体模型结构由 1.3 节给出。该模型用于完成乳腺钼靶肿块检测任务, 训练过程中, 首先使用预训练模型在没有位置标注的钼靶数据集上对主干网络进行训练, 再将主干网络的参数迁移至检测网络的主干网络, 并使用少量有位置标注的钼靶数据对检测网络特有的检测组件进行训练和微调。

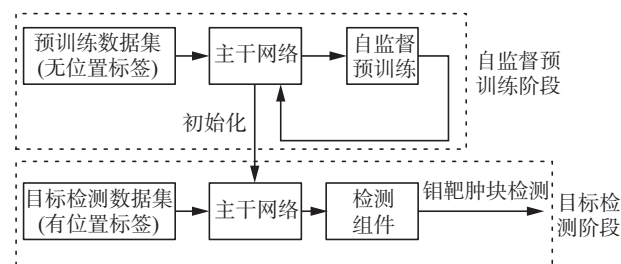


图 1 乳腺钼靶自监督预训练迁移检测模型结构
Fig. 1 Structure of self-supervised pretraining transfer detection model for breast mammographic target

1.2 自监督对比学习预训练模型 EMCL

自监督对比学习通过设置辅助任务, 对比输入的数据来学习有用的表征。不同的对比学习方法往往在正负样本对的构建方法上有所区分, 并且根据不同的辅助任务构建不同的对比损失函数。对比学习按照正负样本对的构建方法分为 3 种。

1) 情境-实例对比学习

利用空间上下文, 通过预测图像补丁相对于其他补丁的位置来学习特征表示。利用上下文信息的算法中, 最流行的算法是最大化互信息 (mutual information, MI)^[14-16]。特征表示的无监督学习可以通过最大化输入图像和由深度神经网络编码的输出图像之间的 MI 来实现。

2) 实例-实例对比学习

在该类别下, 各种方法从 2 个角度进行实例比较。第 1 个是设计或修改对比损失函数, 并使用特定的结构来训练模型^[17-19]。第 2 个角度是直接比较实例, 以获得实例中不同的信息^[20-22]。

3) 时间对比学习

医学成像数据集, 比如 CT (computed tomography) 或 MRI (magnetic resonance imaging) 图像, 包含与扫描相关的一些结构信息。CT 或 MRI 图像序列, 有助于学习更多的语义表示。与二维数据相比, 视频或图像序列具有更丰富的信息, 可以通过对比学习更好地学习特征表示。

由于钼靶图像不涉及空间以及时序信息, 所以实例-实例对比学习方法更适用于钼靶图像的深度学习模型的特征学习。本文受 MoCov2 (momentum contrast)^[23] 的启发, 并在 MoCov2 模型的基础上作改进, 提出基于 EMD 的动量对比学习 (EMD based momentum contrast learning, EMCL)。改进主要包括 2 个方面。

1) 模型结构与损失函数

由于特征提取网络要迁移到检测任务中, 所以图像的空间特征对于下游任务很重要。自监督对比模型中对于相似度的度量往往使用全局池化层生成向量来对比 2 个特征图的相似度, 然而这破坏了空间结构和局部信息。为了保留图像的空间特征和摆脱对全局池化的依赖, 将 MoCov2 中的编码器的全连接层以及映射头去掉, 并在编码器的顶部级联一个残差模块来加强对特征的提取能力, 使用推土距离 (earth mover's distance, EMD) 来度量 2 个特征图的相似度。

EMD 用来表示 2 个分布之间的距离, 已经被许多计算机视觉任务所采用^[24-25]。EMD 衡量了把数据从一种分布移动成另一种分布时需要移动的平均距离的最小值, 其实际上是线性规划中

运输问题的最优解, 应用于 2 个特征图时, 可以将 2 个特征图视为“供应商”和“需求者”, EMD 就是将资源从“供应商”运输到“需求者”所需的最小成本。假设 2 个特征图 $\mathbf{X}, \mathbf{Y} \in \mathbf{R}^{H \times W \times C}$, 将 2 个特征图展开成 $H \times W$ 个 C 维向量的形式, 即 $\mathbf{X} = \{\mathbf{x}_i | i = 1, 2, \dots, H \times W\}$, $\mathbf{Y} = \{\mathbf{y}_j | j = 1, 2, \dots, H \times W\}$, EMD 距离定义公式为

$$d_{\text{EMD}} = \min_{f_{ij}} \frac{\sum_{i=1}^M \sum_{j=1}^N f_{ij} d_{ij}}{\sum_{i=1}^M \sum_{j=1}^N f_{ij}} \quad (1)$$

其中 f_{ij} 需要满足的限定条件为

$$\begin{aligned} \sum_j f_{ij} &\leq x_i \\ \sum_i f_{ij} &\leq y_j \\ \sum_{i=1}^M \sum_{j=1}^N f_{ij} &= \min \left(\sum_i x_i, \sum_j y_j \right) \\ f_{ij} &\geq 0 \end{aligned}$$

式 (1) 中 d_{ij} 表示 \mathbf{x}_i 和 \mathbf{y}_j 之间的距离, 定义为

$$d_{ij} = 1 - \frac{\mathbf{x}_i^T \mathbf{y}_j}{\|\mathbf{x}_i\| \|\mathbf{y}_j\|}$$

通过 EMD 距离求得最佳运输方式 f_{ij} , 则 2 个特征图的相似度得分定义为

$$S(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{H \times W} \sum_{j=1}^{H \times W} f_{ij} (1 - d_{ij})$$

通过 2 个特征图的相似性度量定义自监督对比学习模型的损失函数公式为

$$\ell_{i,j} = -\log \frac{\exp(S(\mathbf{X}, \mathbf{Y})/\tau)}{\sum_{k=1}^{2N} 1_{[k \neq i]} \exp(S(\mathbf{X}, \mathbf{Y})/\tau)}$$

2) 正负样本对构建方法

MoCov2 创建正样本对的方法是将同一图像经过 2 种不同变换方式所得的图像视为一对正样本, 而同一批次以及记忆库中的样本作为负样本。但是由于钼靶图像中良恶性的肿块的差异并不大, 这样创建的正负样本对中正样本与负样本的差异性较小, 容易对模型造成干扰, 所以在这一小节中, 借助类别标签创建新的正负样本对。正样本对的创建方法不变, 在良恶性混淆的数据集中随意选取一个图像作为输入图像, 该图像的 2 个不同视图作为一个正样本对。而负样本则在类别为正常的数据集中抽取, 与类别为良性或者恶性的乳腺钼靶图像作为一个负样本对。这样创建正负样本对是为了加强模型对肿块特征的提取能力, 从而在检测任务中更好地定位肿块位置。EMCL 模型结构如图 2 所示。

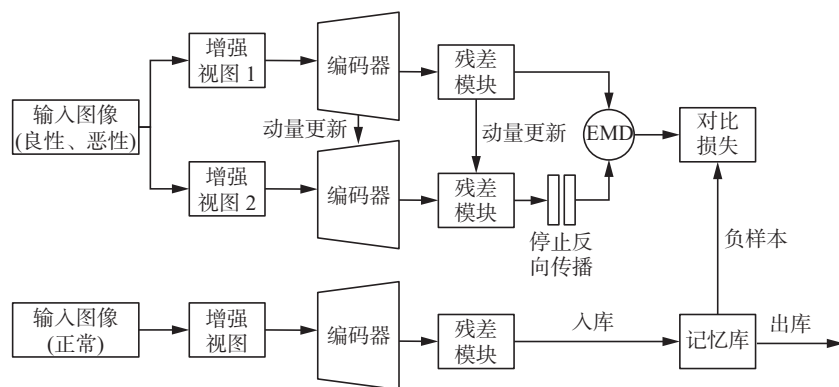


图2 EMCL 模型

Fig. 2 EMCL model

这种改进方法并不是完全不依赖数据的标签的。严格意义上说,该方法在分类任务中不是完全自监督的,但是在钼靶图像检测中只用到了图像的正常或者良恶性的标签,这种标签在临床中可以通过钼靶、超声以及病理切片等其他检查共同诊断得到,与像素级标签或者位置坐标的标签相比,这类标签相对容易获得,在完成检测任务时并不需要具体的位置标签,这种改进方法可以使模型在检测任务中摆脱大部分对标签的依赖。

1.3 乳腺钼靶肿块单阶段目标检测框架设计

现有的特定领域图像目标检测器通常可分为2类,一类是双阶段目标检测器,最具代表性的是 Faster R-CNN (fast region-based convolutional network)^[26]。另一种是单阶段目标检测器,如 YOLO (you only look once)^[27]、FCOS (fully convolutional one-stage object detection)^[28]等。双阶段目标检测器具有较高的定位和目标识别精度,而单阶段目

标检测器具有较快的推理速度。所以本文采用单阶段目标检测器。本文受 FCOS 检测算法的启发,改进了 FCOS 网络架构组成单阶段目标检测器实现乳腺钼靶图像病灶的检测。

由于乳腺钼靶图像属于灰度图像,并不存在颜色特征,并且良恶性的诊断大部分依赖于肿块区域的边缘形状特征,所以更高层的语义特征对于钼靶图像来说是不必要的。特征提取网络中浅层的特征图中保留了更多小目标的信息,而经过下采样后,高层次的特征图更适合对大尺寸的目标进行检测。FCOS 是一种像素级目标检测算法,对于像素数较少的小目标,神经网络的下采样会导致目标信息的丢失,从而使检测模型无法学习到目标的有效特征,进而影响模型的检测精度。本文针对以上问题,对 FCOS 算法的主干网络、FPN(feature pyramid networks)、头部网络分别进行了改进。改进后的单阶段目标检测网络结构如图3所示。

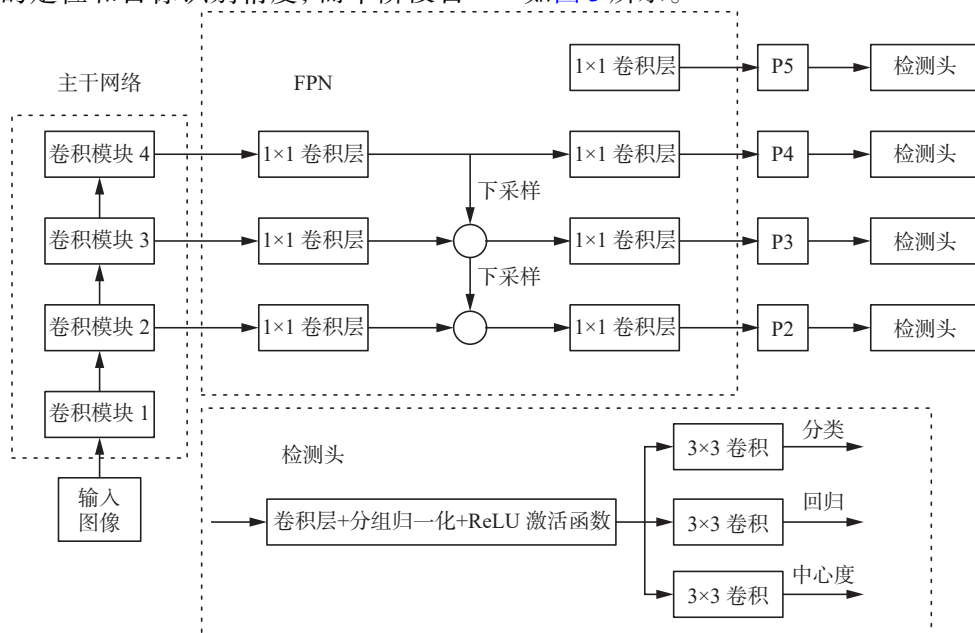


图3 改进的单阶段目标检测网络结构

Fig. 3 Improved one-stage object detection network architecture

改进的细节主要包括 3 点。

1) 主干网络

乳腺钼靶肿块目标的特征信息相对简单,主要的特征信息是模型浅层输出的几何特征。特征信息通常保留在模型浅层的特征层中,FCOS 算法使用的是主干网络中的 C3 (CSP bottleneck with 3 convolutions)、C4 和 C5 特征层,是高层次的特征图,本文在原始 FCOS 网络的基础上,删除了主干网络中的 C5 特征层,增加使用 C2 特征层,即将 C2、C3 和 C4 特征层作为 FPN 的输入特征。这样可以简化模型复杂度,并且可以充分利用图像的低层几何特征,同时也有利于进行对小目标的检测。

2) FPN

FPN 可以增加检测器对于多尺度目标的适应性,下采样后也可以处理 2 个目标框有重叠的情况,但是在乳腺钼靶图像中不存在目标框重叠的情况,所以删除了 FPN 处理后的 P6(C6 的映射)和 P7 特征层,增加了 P2 特征层进行后续的预测,使用经过 FPN 下采样和特征融合后的 P2、P3、P4 和 P5 特征层进行分类、回归和中心度的预测。

3) 检测头

在原有的 FCOS 算法模型中,分类和回归是 2 个独立的分支,在本文中分类和回归分支中的特征提取层共用,这样可以减少网络模型的计算负荷。同时,在这一部分的共享卷积层中,使用可变形卷积网络 (deformable convolution network, DCN) 而不是常规卷积层。

2 数据集来源及处理

本文一共使用到 2 种数据集,分别是乳腺钼靶公开数据集筛查乳腺钼靶数字数据库 DDSM^[29-30]以及来自哈尔滨市肿瘤医院的数据集。

2.1 公共数据集

DDSM 是一个供乳腺钼靶图像分析研究团体使用的资源。该数据库包含大约 2 500 项研究案例。每个研究案例包括 4 张图像,分别为患者的左右乳房的 CC (cranio caudal) 视图和 MLO (mediolateral oblique) 视图,还包括一些相关的患者信息、空间分辨率以及病灶区域的信息描述 (包括类型和位置) 等。病灶区域是由像素级地面实况的遮罩表示的,如图 4 所示,其中图 4(a) 为乳腺钼靶图像的原图,图 4(b) 为对应的 ROI (region of interest) 区域的图像,图 4(c) 为位置信息的像素级标注。

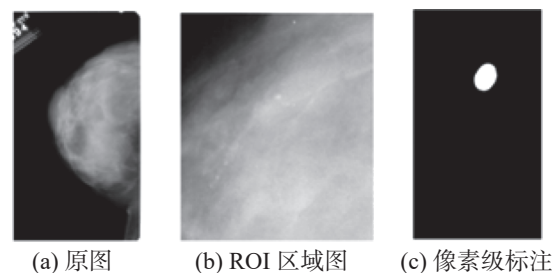


图 4 乳腺钼靶公开数据集 DDSM 数据示例

Fig. 4 Examples of DDSM data from a breast mammographic public dataset

2.2 私有数据集

本文使用的数据集除了公开数据集,还有一部分来源于哈尔滨市肿瘤医院提供的钼靶数据,该数据集共包含 7 520 张图像,每个案例的图像不完整,即同一患者的左右乳房的 CC 和 MLO 的视图并不都包括,数据集的标注是对整张图像的类别标注,缺乏病灶位置和类型的具体标注。整张图的分辨率有 2 082 像素×800 像素和 4 728 像素×5 928 像素 2 种,分别对应图 5(a) 和图 5(b)。

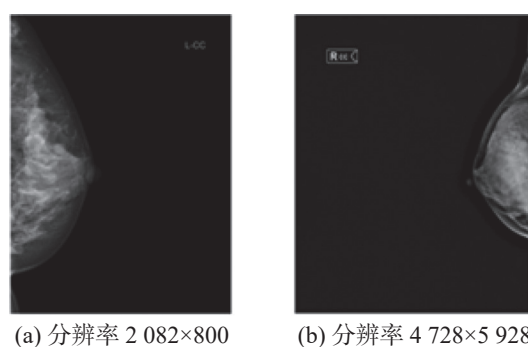


图 5 哈尔滨市肿瘤医院乳腺钼靶数据示例

Fig. 5 Examples of breast mammographic data from Harbin Cancer Hospital

2.3 数据增强

深度学习中,为了增强模型的鲁棒性,常常需要给输入的数据作数据预处理。与计算机视觉任务中通常使用的 RGB 自然图像相比,医学图像的处理面对着特别的挑战。所以在自然图像中应用的数据增强方式并不一定适用于医学图像,于是本文设计了适用于乳腺钼靶图像的数据增强组合策略,包括几何变换和强度操作。

1) 几何变换

几何变换包括随机剪裁、缩放、翻转、随机遮挡、添加噪声等,图 6 给出了几种几何变换对应的数据增强效果。

2) 强度操作

除了几何变换,强度操作也适用于钼靶图像这类的灰度图像,合理的强度操作可以使病灶区域与乳腺组织区域的区别更加明显,常用的

强度操作包括: 调整亮度、调整对比度、调整色相、调整饱和度等。由于钼靶图像是灰度图像, 色彩强度的变化对钼靶图像无效, 其余的强度操作增强方式应用于钼靶图像的效果如图 7 所示。

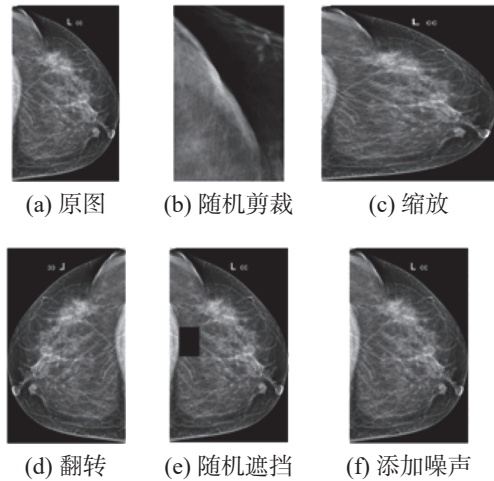


图 6 常用几何变换数据增强效果示例

Fig. 6 Examples of common geometric transformation data enhancement

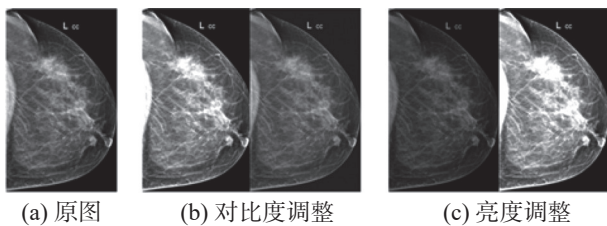


图 7 常用强度操作数据增强效果示例

Fig. 7 Examples of common strength operation data enhancement effects

3) 可变形图像增强

本文也使用了可变形图像增强技术作为数据增强策略之一。为了确保变形之后产生的增强的图片在临床上是可信的, 变形的规模和程度可以由参数限制在一定范围内。应用可变形增强技术还可以模拟在临床扫描中可能遇到的一系列合理的变化, 例如组织变形, 或病人移动产生的图像伪影。可变形的数据增强技术使用随机位移场(random displacement fields, RDF) 的方法, 其本质是一种仿射变换和弹性变形。使用该方法进行数据增强需要 3 个步骤。

1) 将图像 A 进行三点法变换得到图像 B。仿射变换是一种二维坐标到二维坐标之间的线性变换, 可以保持图形的平直性。

2) 对于图像 B 中每个像素点随机生成一个在 x 和 y 方向的位移, 分别记为 Δx 和 Δy 。位移范围控制在 $(-1, 1)$ 区间内, 这样就得到了一个随机

位移场。

3) 用服从高斯分布 $N(0, \sigma)$ 的高斯函数对步骤 2) 中产生的随机位移场进行滤波操作, 其中高斯分布的标准差 σ 影响着图像的平滑程度, σ 越大, 产生的图像越平滑。

对图像进行随机位移场操作后的效果如图 8 所示。

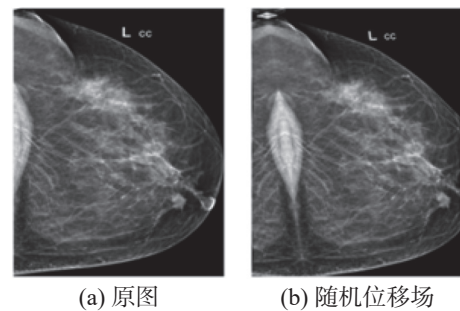


图 8 可变形数据增强变换效果

Fig. 8 Deformable data enhances the transformation effect

3 实验与分析

本实验的编程环境和平台搭建是在 Ubuntu 16.04 操作系统下完成的, GPU 选用 NVIDIA GeForce RTX 3060, 显存大小为 12 GB, CUDA 版本为 11.2, 编程语言采用 Python 3.8。在对数据集进行数据增强处理之后, 将良性和恶性数据分别按照训练集: 验证集: 测试集=8:1:1 进行划分, 用于对模型进行训练、验证和测试。

3.1 训练细节

首先使用 ResNet50^[31] 作为主干网络, 图像分类作为下游任务完成数据增强策略组合的有效性验证实验, 实验采用 ImageNet^[32] 预训练迁移学习来加速网络收敛, 并分别在公共数据集与私有数据集上进行测试, 记为实验 1。然后在私有数据集使用本文改进的 MoCov2 模型进行主干网络自监督训练, 主干网络使用 ResNet50, 目的是与检测模型的特征提取网络保持一致。这一部分称为预训练阶段, 且沿用了实验 1 中筛选的最有效的数据增强策略, 记为实验 2。最后将训练好的主干网络迁移至改进的 FCOS 检测模型中, 冻结主干网络 C2、C3 层参数, 使用公共数据集进行 C4、FPN 以及头部网络的训练, 这一部分称为微调阶段, 记为实验 3。实验 1~3 的超参数设置分别如表 1~3 所示。实验 1 使用准确率 (accuracy, ACC) 作为评价指标来验证数据增强的有效性。实验 2、3 的检测任务使用平均精度 (average precision, AP) 与均值平均精度 (mean average precision, mAP) 来评价检测算法在数据集上的测试结果。

表 1 数据增强策略筛选实验超参数设置 (实验 1)

Table 1 Data enhancement strategy screening experiment super parameter setting (experiment 1)

优化器相关参数				批量大小	迭代次数
类别	学习率	冲量	权重衰减		
16	100	SGD	0.001	0.9	0.000 1

表 2 预训练阶段超参数设置 (实验 2)

Table 2 Super parameter setting in pre-training stage (experiment 2)

优化器相关参数				批量大小	代次数	内存库容量	动量更新超参数
类别	学习率	冲量	权重衰减				
16	100	16 384	0.999	SGD	0.015	0.9	0.000 4

表 3 微调阶段超参数设置 (实验 3)

Table 3 Superparameter setting in fine tuning stage (experiment 3)

优化器相关参数				批量大小	迭代次数
类别	学习率	冲量	权重衰减		
8	200	SGD	0.005	0.9	0.000 1

3.2 实验结果及分析

3.2.1 数据增强方式组合测试

不同数据增强方式对特征提取网络的影响如表 4 所示。由于卷积神经网络需要保证输入图片的大小一致,所以在随机剪裁和缩放 2 种操作中选择 1 种。对比数据可以看出缩放的效果要优于随机剪裁,这是因为随机剪裁可能会对图像中关键信息区域造成破坏,得到伪影或者

背景区域,对下游任务造成干扰。旋转和翻转都是有效的数据增强方式,加入后模型的性能有所提升。而加入随机遮挡后性能有所下降,证明了随机遮挡也有可能造成关键信息丢失的问题。加入了可变形图像增强后模型效果再一次提升,所以得出最有效的数据增强策略组合为:缩放、旋转、翻转、强度调整以及可变形图像增强。

表 4 不同数据增强方式组合的性能差异

Table 4 Performance of different combinations of data augment methods

随机裁剪	翻转	旋转	缩放	随机遮挡	强度调整	可变形技术	公共数据集	私有数据集
							ACC	ACC
√			√				0.795 8	0.763 9
		√	√				0.794 7	0.770 3
	√	√	√				0.800 3	0.772 1
	√	√	√				0.801 8	0.780 9
	√	√	√		√		0.806 5	0.784 3
	√	√	√		√	√	0.817 2	0.810 3
	√	√	√	√			0.770 5	0.756 9
	√	√	√		√		0.751 2	0.737 0

3.2.2 预训练迁移有效性验证

为了验证钼靶图像自监督预训练迁移至检测网络的方法是有效的,将 ImageNet 预训练迁移、随机初始化、自监督预训练迁移的 3 种主干网络放在改进的检测网络中进行测试。其中自监督预训练模型 EMCL 对比了 SimCLR^[20]、BY-

OL^[33]、MoCov2^[23]。实验结果由表 5 给出。实验结果表明,EMCL 模型在目标检测任务中取得了最优的效果,证明了相比于其他常用于分类任务中的自监督预训练模型,改进的模型可以保留图像的空间特征,对下游检测任务具有积极的影响。

表 5 预训练迁移方法性能对比

Table 5 Performance comparison of pre-training transfer methods

评价指标	随机初始化	ImageNet迁移	SimCLR	BYOL	MoCov2	EMCL
良性AP	0.769 3	0.768 7	0.759 7	0.712 3	0.769 0	0.784 6
恶性AP	0.779 5	0.780 4	0.770 9	0.719 1	0.779 2	0.790 5
mAP	0.774 4	0.774 6	0.765 3	0.715 7	0.774 1	0.787 6

3.2.3 FCOS 检测模型对比

为了验证本文改进的 FCOS 检测模型对于钼靶图像是有效的, 使用公共数据集在多种检测算法中进行实验, 对比的方法包括: Faster R-CNN^[26]、YOLOv5^[27] 以及 FCOS^[28], 以上 3 种算法是医学图像领域使用较为广泛的检测算法。实验结果由图 9~11 给出。实验结果表明, Faster R-CNN 检测

器在良恶性检测和综合检测性能表现较差, 且其训练过程复杂, 检测时间长。在 YOLOv5 和 FCOS 中, 良性肿块检测的 AP 更高的是 FCOS 算法; 恶性肿块检测 YOLOv5 算法要更高一些。4 种算法中, 改进的 FCOS 在良恶性上的 mAP 最高, 同时从 PR 曲线上看, 改进的 FCOS 的曲线更靠近右上方, 说明改进的 FCOS 的综合性能更好。

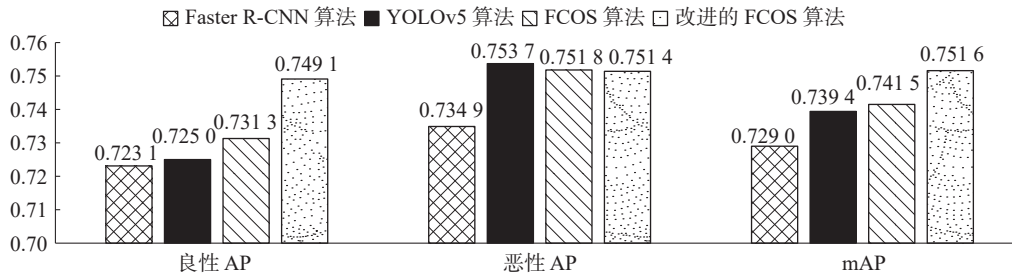


图 9 改进的检测模型与其他检测模型对比

Fig. 9 The improved detection model compared with other detection models

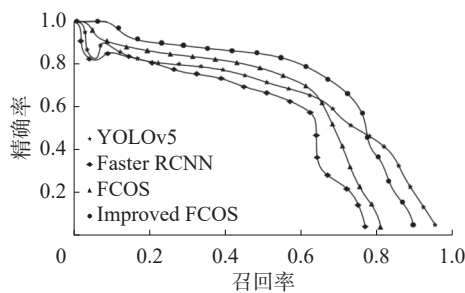


图 10 4 种算法在良性样本上的 PR 曲线

Fig. 10 PR curves of four algorithms on benign samples

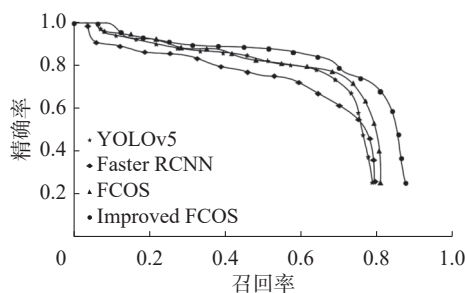


图 11 4 种算法在恶性样本上的 PR 曲线

Fig. 11 PR curves of four algorithms on malignant samples

4 结束语

本文提出了一种自监督对比学习方法 EM-CL 算法, 算法保留了图像的空间特征, 实验表明使用 EMCL 算法学习到的特征表示相比于 SimCLR、BYOL、MoCov2 等自监督对比学习算法更适用于检测的下游任务。在乳腺钼靶肿块检测的任务中, 改进的 FCOS 检测模型简化了原有的结构, 充分利用乳腺钼靶图像的低层几何特征, 使乳腺钼靶肿块的检测效果有所提升, 尤其是良性肿块的检测效果提升更为明显。而良性肿块的尺

寸大多较小, 证明了改进的网络提高了对小目标的检测能力。在后续的探索中将进一步完善基于 EMCL 预训练的检测算法对于其他医学图像的检测性能, 例如 CT、肾脏以及乳腺超声等与自然图像有着明显区别的医学图像, 并有望在医学图像领域生成一个代替 ImageNet 的更有效且更适用于目标检测的预训练迁移模型。

参考文献:

- [1] 文帅, 李华兵. 放射学及放射组学在骨质疏松症的研究进展 [J]. 中国 CT 和 MRI 杂志, 2023, 21(9): 181-183.
WEN Shuai, LI Huabing. Research progress of radiology and radiomics in osteoporosis [J]. Chinese journal of CT and MRI, 2023, 21(9): 181-183.
- [2] 刘姣. 放射学检查方法在肿瘤诊断中的研究进展 [J]. 中国医药指南, 2023, 21(7): 67-69.
LIU Jiao. Research progress of radiological examination methods in tumor diagnosis [J]. Guide of China medicine, 2023, 21(7): 67-69.
- [3] 张政. 人工智能在皮肤病诊断中的应用 [J]. 中国新通信, 2020, 22(24): 118-119.
ZHANG Zheng. Application of artificial intelligence in the diagnosis of dermatosis [J]. China new telecommunications, 2020, 22(24): 118-119.
- [4] 于鹤, 赵稳兴. 计算机辅助诊断技术在病理学中的应用进展 [J]. 诊断病理学杂志, 2018, 25(3): 223-226.
YU He, ZHAO Wenxing. Application progress of computer-aided diagnosis technology in pathology [J]. Chinese journal of diagnostic pathology, 2018, 25(3): 223-226.
- [5] 徐立芳, 傅智杰, 莫宏伟. 基于改进的 YOLOv3 算法的乳腺超声肿瘤识别 [J]. 智能系统学报, 2021, 16(1):

- 21–29.
- XU Lifang, FU Zhijie, MO Hongwei. Tumor recognition in breast ultrasound images based on an improved YOLOv3 algorithm[J]. *CAAI transactions on intelligent systems*, 2021, 16(1): 21–29.
- [6] DHUNGEL N, CARNEIRO G, BRADLEY A P. The automated learning of deep features for breast mass classification from mammograms[C]//Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2016: 106–114.
- [7] KRITI, VIRMANI J, DEY N, et al. PCA-PNN and PCA-SVM based CAD systems for breast density classification[M]//Intelligent Systems Reference Library. Cham: Springer, 2015: 159–180.
- [8] KETABI H, EKHLASI A, AHMADI H. A computer-aided approach for automatic detection of breast masses in digital mammogram via spectral clustering and support vector machine[J]. *Physical and engineering sciences in medicine*, 2021, 44(1): 277–290.
- [9] SANNASI CHAKRAVARTHY S R, RAJAGURU H. Automatic detection and classification of mammograms using improved extreme learning machine with deep learning[J]. *Innovation and research in biomedical engineering*, 2022, 43(1): 49–61.
- [10] GNANASEKARAN V S, JOYPAUL S, MEENAKSHI SUNDARAM P, et al. Deep learning algorithm for breast masses classification in mammograms[J]. *IET image processing*, 2020, 14(12): 2860–2868.
- [11] AL-MASNI M A, AL-ANTARI M A, PARK J M, et al. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system[J]. *Computer methods and programs in biomedicine*, 2018, 157: 85–94.
- [12] 韩昌芝, 俞璐, 李林, 等. 无源无监督迁移学习综述 [J]. *信息技术与信息化*, 2023(12): 46–51.
- HAN Changzhi, YU Lu, LI Lin, et al. A summary of passive unsupervised transfer learning[J]. *Information technology and informatization*, 2023(12): 46–51.
- [13] 邓佳丽, 龚海刚, 刘明. 基于目标检测的医学影像分割算法 [J]. *电子科技大学学报*, 2023, 52(2): 254–262.
- DENG Jiali, GONG Haigang, LIU Ming. Medical image segmentation based on object detection[J]. *Journal of university of electronic science and technology of China*, 2023, 52(2): 254–262.
- [14] WANG W C, AHN E, FENG Dagan, et al. A review of predictive and contrastive self-supervised learning for medical images[J]. *Machine intelligence research*, 2023, 20(4): 483–513.
- [15] TSCHANNEN M, DJOLONGA J, RUBENSTEIN P K, et al. On mutual information maximization for representation learning[EB/OL]. (2019–07–31)[2023–04–15]. <http://arxiv.org/abs/1907.13625>.
- [16] HJELM R D, FEDOROV A, LAVOIE-MARCHILDON S, et al. Learning deep representations by mutual information estimation and maximization[EB/OL]. (2018–08–20)[2023–04–15]. <http://arxiv.org/abs/1808.06670>.
- [17] TIAN Yonglong, KRISHNAN D, ISOLA P. Contrastive multiview coding[C]//European Conference on Computer Vision. Cham: Springer, 2020: 776–794.
- [18] AGRAWAL P, CARREIRA J, MALIK J. Learning to see by moving[C]//2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 37–45.
- [19] CHEN Xinlei, HE Kaiming. Exploring simple Siamese representation learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 15745–15753.
- [20] CHEN Ting, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations[C]//Proceedings of the 37th International Conference on Machine Learning. Vienna: PMLR, 2020: 1597–1607.
- [21] HE Kaiming, FAN Haoqi, WU Yuxin, et al. Momentum contrast for unsupervised visual representation learning [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 9726–9735.
- [22] GRILL J B, STRUB F, ALTCHÉ F, et al. Bootstrap your own latent: a new approach to self-supervised Learning [EB/OL]. (2020–06–13)[2023–04–15]. <http://arxiv.org/abs/2006.07733>.
- [23] CHEN Xinlei, FAN Haoqi, GIRSHICK R, et al. Improved baselines with momentum contrastive learning [EB/OL]. (2020–03–09)[2023–04–15]. <http://arxiv.org/abs/2003.04297>.
- [24] WANG C, CHAN S C. A new hand gesture recognition algorithm based on joint color-depth Superpixel Earth Mover’s Distance[C]//2014 4th International Workshop on Cognitive Information Processing. Copenhagen: IEEE, 2014: 1–6.
- [25] NIKOLENTZOS G, MELADIANOS P, VAZIRGIAN-NIS M. Matching node embeddings for graph similarity [C]//Proceedings of the AAAI Conference on Artificial Intelligence. San Francisco: AAAI, 2017: 2429–2435.
- [26] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(6): 1137–1149.
- [27] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern

Recognition. Las Vegas: IEEE, 2016: 779–788.

- [28] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: fully convolutional one-stage object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 9626–9635.
- [29] HEATH M, BOWYER K, KOPANS D, et al. The digital database for screening mammography[M]. Nijmegen: Springer, 2001: 212–218.
- [30] HEATH M, BOWYER K, KOPANS D, et al. Current status of the digital database for screening mammography[M]. Dordrecht: Springer, 1998: 457–460.
- [31] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770–778.
- [32] FEI-FEI L, DENG J, LI K. ImageNet: Constructing a large-scale image database[J]. *Journal of vision*, 2010, 9(8): 1037.
- [33] GRILL J B, STRUB F, ALTCHÉ F, et al. Bootstrap your own latent: a new approach to self-supervised learning[EB/OL]. (2020–06–13)[2023–04–15]. <https://arxiv.org/abs/2006.07733>.

作者简介:



莫宏伟, 教授, 博士生导师, 主要研究方向为类脑计算与人工智能、机器视觉与机器认知、人机混合智能。主持省部级科研项目 24 项, 授权发明专利 10 项, 发表学术论文 80 余篇。
E-mail: honwei2004@126.com。



孙琪, 硕士研究生, 主要研究方向为深度学习、计算机视觉、医学影像。
E-mail: 962287826@qq.com。



孙鹏, 硕士研究生, 主要研究方向为深度学习、计算机视觉、医学影像。
E-mail: 1940737851@qq.com。

2024 中国智能机器人生态大会

2024 年 9 月 8 日, 全国科普日来临之际, 由中国人工智能学会(CAAI)、合肥市人民政府、安徽省人工智能产业推进组办公室(安徽省科学技术厅)主办, CAAI 智能决策专委会、江淮前沿技术协同创新中心、合肥市智能机器人产业专班办公室(合肥市发展和改革委员会)、安徽中安创谷科技园有限公司联合承办的中国智能机器人生态大会将在合肥举办。中国人工智能学会智能决策专委会成立大会同期举行。

本次会议主题为“智能科技 决策未来”, 将邀请人工智能和机器人领域多位院士专家及具有代表性的科研院所、高新技术企业、投资机构参加。大会聚集智能机器人领域前沿趋势, 围绕决策大模型、多智能体强化学习、人形机器人和产业发展等热点议题进行主题分享和对话交流, 旨在搭建人工智能学术交流合作平台, 构建智能机器人产业创新生态, 助力我国智能机器人产业高质量发展。

主题报告:

戴琼海 中国工程院院士 《人工智能发展与思考》
樊邦奎 中国工程院院士 《浅谈低空经济》
郭 雷 中国科学院院士 《无人系统仿生智能——理论与实践探索》
安徽省“AI+机器人”科创发展座谈

主题论坛:

决策大模型论坛
多智能体强化学习论坛
人形机器人技术创新论坛
智能机器人产业协同发展论坛

会议期间将推介安徽省通用人工智能专项政策、合肥市智能机器人产业行动计划、合肥市 2024 年智能机器人领域“揭榜挂帅”项目榜单、合肥市智能机器人“超级场景”等。会议现场还将发布江淮前沿技术协同创新中心自研人形机器人“启江二号”。