

基于 Vision Transformer 的深部隐伏矿体 三维成矿预测方法研究

吴以婕^{1,2,3}, 李晓晖^{1,2,3}, 袁峰^{1,2,3}, 郑超杰^{1,2,3}, 徐艳^{1,2,3}, 张明明^{1,2,3}

1. 合肥工业大学资源与环境工程学院, 安徽 合肥 230009
2. 安徽省战略性矿产资源深部探测与评价利用重点实验室, 安徽 合肥 230009
3. 安徽省矿产资源与矿山环境工程技术研究中心, 安徽合肥 230009

摘要: 三维成矿预测是深部隐伏矿产资源勘查重要的方法技术之一。近年来, 以卷积神经网络为代表的深度学习方法在三维成矿预测信息融合方面取得一定研究进展, 但受限于卷积神经网络的局部感受野, 可能难以提取三维预测要素与矿化事实之间的长程依赖与全局关联, 制约了深部隐伏矿体的预测精度。针对上述问题, 本研究基于 Vision Transformer (ViT) 架构, 构建了适用于三维地质体数据的 3D-ViT 模型。模型通过 3D 体素块嵌入模块和分离式三维位置编码, 显式保留地质体的结构信息, 借助多头自注意力机制构建全局感知场, 以期对岩体、地层、构造等多预测要素建立与矿化事实的跨尺度的空间关联进行建模。在安徽省狮子山矿田的实例研究中, 该模型成功预测了主要已知矿体, AUC 值达到 0.96, 其准确率、召回率与 F1 分数均优于 3D-CNN 及传统机器学习模型, 展现出良好的预测能力和预测精度。基于预测结果, 研究最终在狮子山矿田深部圈定了 4 处找矿靶区, 验证了该方法在复杂地质结构下捕捉隐蔽矿化信息的有效性与可靠性。本研究不仅拓展了 ViT 在地学三维数据中的应用范畴, 也为深部矿产资源智能预测提供了具有全局感知能力的新方法论支持, 具备重要的勘查应用前景。

关键词: 三维成矿预测; 深部隐伏矿体; 视觉 Transformer (ViT)

中图分类号: P612 收稿日期: 2025-11-29

Vision Transformer based 3D Mineral Prospectivity modeling for Deep Concealed Ore Bodies

Wu Yijie^{1,2,3}, Li Xiaohui^{1,2,3}, Yuan Feng^{1,2,3}, Zheng Chaojie^{1,2,3}, Xu Yan^{1,2,3}, Zhang Mingming^{1,2,3}

1. School of Resources and Environmental Engineering, Hefei University of Technology, Hefei 230009, China
2. Anhui Provincial Key Laboratory for Deep Exploration, Evaluation and Utilization of Strategic Mineral Resources, Hefei 230009, China
3. Anhui Province Engineering Research Center for Mineral Resources and Mine Environments, Hefei 230009

Abstract: Three-dimensional mineral prospectivity modeling serves as a crucial technical approach in the exploration of deep concealed mineral resources. In recent years, deep learning methods represented by convolutional neural networks have achieved some progress in integrating 3D predictive information; however, constrained by the local receptive fields of CNNs, it remains difficult to extract long-range dependencies and global correlations between 3D predictive factors

基金项目: 国家深地重大科技专项(No.2025ZD1007402), 国家自然科学基金项目(No.42230802, 42472359)

作者简介: 吴以婕 (2002-), 硕士研究生, 主要从事三维 GIS 与成矿预测方向研究。ORCID:

0009-0001-5853-1328

通讯作者简介: 李晓晖 (1986-), 副教授, 博导, 主要从事三维 GIS 与数学地质学方向研究工作。E-mail:

lxhlixiaohui@163.com

and mineralization occurrences, which limits the prediction accuracy for deep concealed ore bodies. To address these issues, this study develops a 3D-ViT model based on the Vision Transformer (ViT) architecture, tailored for 3D geological data. The model employs a 3D voxel-patch embedding module and decoupled 3D positional encoding to explicitly preserve the structural information of geological bodies. By leveraging a multi-head self-attention mechanism, a global perceptual field is constructed to model cross-scale spatial relationships between multiple predictive factors—such as intrusions, strata, and structures—and mineralization evidence. In a case study of the Shizishan ore field in Anhui Province, the model successfully predicted the main known ore bodies, achieving an AUC of 0.96. Its accuracy, recall, and F1-score all surpassed those of 3D-CNN and traditional machine learning models, demonstrating strong predictive capability and precision. Based on the prediction results, four prospective target areas were delineated in the deep part of the Shizishan ore field, verifying the method's effectiveness and reliability in detecting concealed mineralization under complex geological settings. This research not only extends the application of ViT to three-dimensional geoscientific data but also provides a novel methodology with global perception for intelligent prediction of deep mineral resources, holding significant potential for practical exploration applications.

Keywords: 3D Mineral Prospectivity modeling; Deep Concealed Ore Bodies; Vision Transformer (ViT)

引言

矿产资源勘查开发是国家经济安全和战略发展的重要保障 (Singer, 2018; 翟明国等, 2019)。目前, 中国东部地区浅部易识别矿床已日趋枯竭, 矿产勘查的重点已逐步从浅层矿床转向深部隐伏矿床。然而, 隐伏矿床由于其埋藏深度大以及成矿过程复杂等特点, 使得其定位和预测非常具有挑战性 (Li et al., 2015; Wang et al., 2015; 毛先成等, 2016; Deng et al., 2022)。这使得矿产勘查工作难度激增, 对高精度、智能化的预测理论与技术提出了前所未有的迫切需求。

近年来, 机器学习及深度学习技术飞速发展, 正在推动矿产勘查预测方法从传统的定性类比和统计模型, 向数据驱动的智能算法时代跨越 (Zuo et al., 2025)。这些智能算法可以有效地集成多源异质数据, 如地质、地球物理、地球化学和遥感数据, 以构建高精度的预测模型, 展现出超越传统方法的预测性能 (Yousefi et al., 2024)。机器学习在矿产预测中的发展, 清晰地展现了一条从简单模型到复杂模型的技术演进路径。一系列机器学习模型, 从早期的单一线性模型, 发展到复杂的集成学习与深度学习模型, 已在矿产资源预测中证明了其研究价值 (Sun et al., 2020; Xu et al., 2021; Mao et al., 2024; Xiao et al., 2025)。其中, 深度学习因其强大的非线性建模与特征提取能力, 在矿物勘探领域受到广泛关注, 在成矿预测中展现出巨大潜力 (Xu et al., 2025)。应用最广泛的技术之一是卷积神经网络 (CNN), 它在模式识别和图像处理方面表现出了优越的性能 (Krizhevsky et al., 2017)。目前, CNN 凭借其在空间特征提取方面的优势, 在二维和三维矿产预测中取得了显著成效 (Zhang et al., 2021; Yang and Zuo, 2024)。在以往研究中, 早期侧重于将 CNN 应用于二维地质空间, 如 Li et al. (2020) 结合卷积神经网络算法和迁移学习方法, 并结合地球化学数据, 准确地识别矿化区域; 之后拓展至三维地质空间, 实现了对深部三维地质结构和预测信息空间的有效解读 (Li et al., 2023); 随着研究的加深, Li et al. (2024) 在此之后提出一种融合空间注意力机制和 Inception 模块的多尺度三维卷积神经网络模型 (MSAM-CNN), 对关键预测信息

的捕捉能力得到进一步提升，在三维成矿预测中取得显著成效。尽管 CNN 在成矿预测领域方面取得了较多进展，但由于其基于窗口学习特征向量，专注于短距离局部依赖信息的计算，严重依赖于卷积核在局部邻域内的操作，难以有效捕捉地质控矿因素在区域尺度上的长距离依赖关系与全局空间关联（Wang et al., 2023），因此其预测结果难以融合全局视角下融合由远程预测要素提供的关键信息。

针对智能成矿预测面临的挑战，Zuo（2025）提出了地质约束的自监督图-Transformer 模型，该模型通过图结构捕捉地质实体间的空间耦合关系，并利用 Transformer 架构增强对长程空间依赖的建模能力，为三维地质空间中全局特征的提取提供了新思路。Transformer 架构（Vaswani et al., 2017）的出现为处理长距离依赖和全局特征建模提供了新思路。其在图像识别、自然语言处理等领域的成功应用显示了其在捕捉全局上下文信息方面的强大能力（Fu et al., 2023）。Vision Transformer（ViT）作为其在计算机视觉领域的成功拓展（Dosovitskiy, 2020），已在多个任务中展现出超越 CNN 的潜力，视觉大模型则擅长空间信息提取，为地学数据中全局空间特征的提取提供了新方法（Ning et al., 2025）。目前，Transformer 架构因其强大的全局上下文建模能力，已被引入成矿预测领域。其核心思路在于利用自注意力机制捕捉地质空间中的长程依赖，通过将 Transformer 作为特征增强模块，与 CNN（Li et al., 2024）或图神经网络（Gao et al., 2025）相结合，构建混合模型，以耦合局部特征与全局语义；或者是探索 ViT 框架的适用性，尝试无监督学习策略（Ning et al., 2025）。这些研究共同昭示了全局感知模型在未来智能矿产预测中的巨大潜力。

但是，上述探索均局限应用于二维成矿预测研究工作，ViT 在三维地质空间中的成矿预测研究尚未有效开展，其应用于三维体数据时的模型适应性与预测效能亟待深入探索与验证。在此背景下，本文依托“四步式”三维成矿预测方法（袁峰等，2018），提出一种基于 Vision Transformer 的三维成矿预测方法，并在狮子山矿田开展实际的三维成矿预测研究，以验证模型方法在深部找矿预测中的可行性与有效性。

1 方法原理

1.1 Vision Transformer 核心原理

Vision Transformer（ViT）是将 Transformer 架构应用于视觉领域的典型模型，由 Dosovitskiy（2020）提出，是近年来计算机视觉领域的重要突破。ViT 的核心思想是将输入图像分割为固定大小的图像块（Patch），从而将二维图像转换为序列化的数据，使 Transformer 架构能够直接用于图像处理任务。具体而言，这些图像块通过线性投影转换为嵌入向量，并添加位置编码以保留其空间信息；随后，该嵌入序列被输入到由多层组成的标准 Transformer 编码器中，通过其核心的自注意力机制（Vaswani et al., 2017）建立全局依赖关系。与 CNN 依赖局部卷积核逐层传递信息相比，ViT 能从初始阶段就通过自注意力机制捕获空间中的长程依赖关系，从而建立全局上下文关联，并且其特征映射能更有效地保留输入的空间位置信息。

1.2 模型整体架构设计

针对三维成矿预测的特殊需求，本研究在标准的 ViT 架构基础上进行了改造，构建适用于三维成矿预测场景的 ViT 模型。模型采用体素块（Voxel Patch）作为基本输入单元，通过三维位置编码与多头自注意力结构实现空间特征学习，为深部隐伏矿体预测提供了新的技术途径。模型整体架构如图 1 所示，主要由输入层、特征嵌入层、Transformer 编码器和分类输出层组成。输入层中 3D Patch Embedding 模块将输入数据转换为序列化的令牌表示；随

后，添加可学习的位置编码以保留三维空间结构信息，在 Transformer 编码器中利用自注意力机制进行深度特征提取后；最后在分类输出层通过分类头输出成矿有利度。

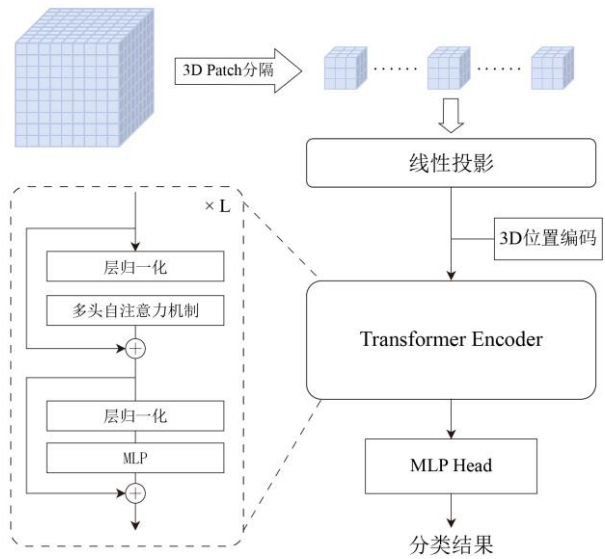


图 1 Vision Transformer 模型架构图

Figure 1 Vision Transformer (ViT) model architecture diagram

1.2.1 3D Patch Embedding 模块

与原始 ViT 处理二维图像不同，本研究的 3D Patch Embedding 专门针对三维地质数据。模型的输入为规则化的三维地质数据块，采用滑动窗口法，以每个网格单元为中心提取一个固定尺寸的局部三维数据块作为一次输入的样本，其空间维度（长×宽×高）可根据研究区实际网格大小灵活定义，特征通道数则对应于所筛选的成矿有利要素数量。为适配 Transformer 架构，模型首先通过 3D Patch 划分，将数据块在空间上划分为多个不重叠的小立方体（Patch），再进行嵌入和序列化处理。每个 Patch 通过一个线性投影层被映射到一个高维的嵌入空间，从而将原始地质特征转换为一系列可供 Transformer 编码器处理的序列化令牌。这一处理流程将局部空间域的地质特征转化为特征域的可学习向量，为后续的全局自注意力计算奠定了基础。

与直接将整个地质体展平相比，Patch 划分显著降低了计算复杂度，使模型训练更加可行。除此之外，较小的 Patch 尺寸能够捕获局部地质特征，而通过自注意力机制的全局交互，模型能够整合这些局部信息，理解区域成矿系统的整体结构。

1.2.2 3D 位置编码机制

位置编码是赋予 Transformer 模型空间感知能力的关键。原始 ViT 采用一维可学习位置编码或正弦编码，这些方法在三维空间中存在明显不足。隐伏矿体预测需要精确表示三维空间中体素的结构位置，因此模型采用分离式三维位置编码，分别为深度、高度和宽度三个空间维度学习独立的可学习位置编码，并在嵌入阶段将它们相加融合，使模型能够感知三维空间结构及位置关系。

这种分离式设计能够更好地建模三维空间中的各向异性特征，符合地质体在空间中通常表现出的方向性差异，区分不同方向上的地质趋势差异，使模型能够更好地理解三维地质环境中的成矿控制因素。确保模型在处理每个位置时，既考虑其地质特征，又考虑其空间位置。

1.2.3 Transformer 编码器与多头自注意力机制

此模块是模型的核心处理单元，由 L 层相同的 Transformer 编码器堆叠而成。每一层都

包含一个多头自注意力模块和一个前馈网络,并均采用残差连接与层归一化来保障训练的稳定性。

多头自注意力机制通过线性变换生成查询矩阵 Q 、键矩阵 K 和值矩阵 V , 使序列中的任何位置都能与所有其他位置直接交互, 其公式为:

$$\text{Attention}(Q,K,V)=\text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

式中 d_k 为键向量维度, 缩放因子 $\sqrt{d_k}$ 用于稳定梯度传播。模型将这种自注意力运算在多个头上并行执行, 使模型能够从不同的表示子空间联合关注来自不同位置的信息。在三维成矿预测中, 不同的注意力头可以专注于学习不同类型的地质关联。通过多个注意力头的并行计算, 模型能够有效捕捉长程的地质依赖关系, 从而实现对复杂成矿系统的多视角、全局性理解, 实现了全局感知能力。

为了进一步提高模型的泛化能力, 在每个 Transformer 块中引入了随机深度(DropPath)正则化。在训练过程中, 以一定概率随机丢弃整个子层, 从而鼓励网络学习更鲁棒的特征。丢弃概率随着网络深度线性增加, 即较深的块有更高的概率被丢弃。这种渐进式丢弃策略在保持模型容量的同时, 有效防止了过拟合现象。

2 实例研究区地质概况

2.1 铜陵矿集区地质概况

铜陵矿集区地处长江中下游多金属成矿带, 是我国重要的铜、铁、金多金属矿集区(图2)。在大地构造上, 其处于大别造山带与江南过渡带之间的下扬子拗陷带中部的铜陵凸起内, 隶属于下扬子地层分区(翟裕生等, 1992)。该区域深部构造特征显著, 壳幔分层具有明显的层状结构, 整体位于地幔隆起带上, 地壳内部则呈现出典型的“三层结构”, 这种独特的深部构造背景为成矿作用提供了有利条件。

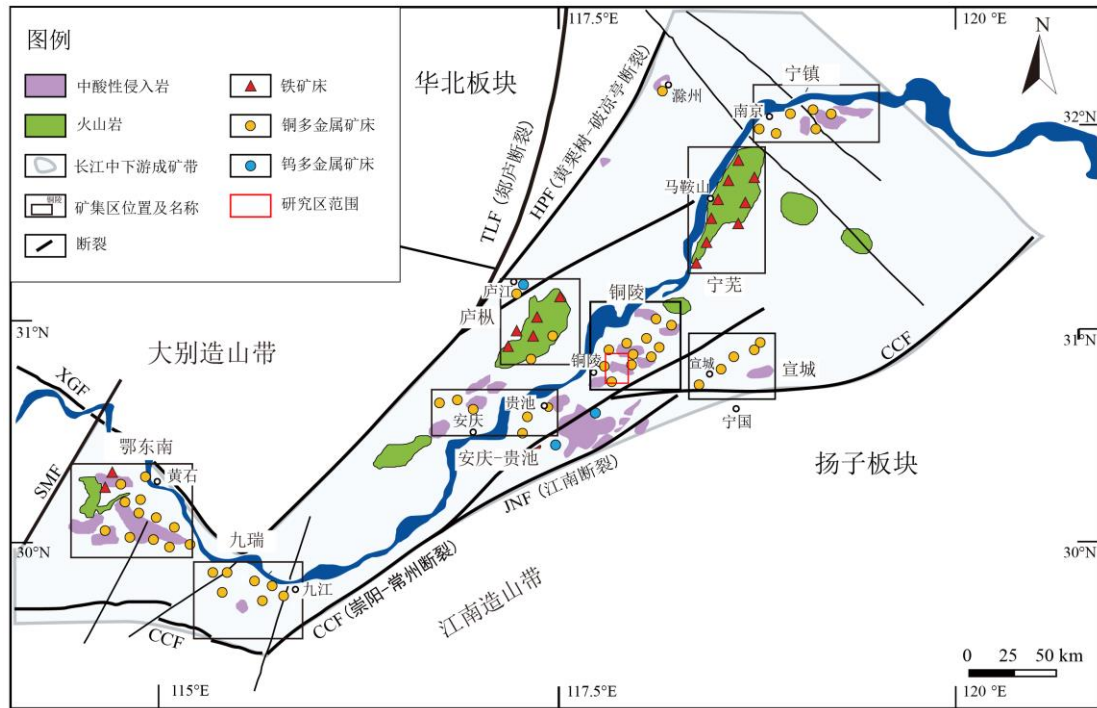


图2 长江中下游成矿带区域地质简图(周涛发等, 2017)

Figure 2 Simplified regional geological map of the Middle-Lower Yangtze River Metallogenic Belt (Zhou et al., 2017)

在地层方面，区内出露地层从志留系直至第四系均有分布，其中以志留系至第三系地层最为发育，矿床的形成与特定地层关系密切，其矿体大多赋存于石炭纪、二叠纪和三叠纪碳酸盐地层之中（毛景文等，2009），其中石炭系的黄龙组、船山组以及三叠系下统的南陵湖组构成了区内最主要的成矿层位。

铜陵矿集区的地质构造格局复杂，岩浆活动十分强烈，主要发生在燕山期至白垩纪，并生成了系列具有重要成矿意义的中酸性侵入岩，主要包括辉石闪长岩、石英闪长岩、石英二长岩及花岗闪长岩类。这些岩体的空间展布规律，主要受北北东向与近东西向构造的复合作用所控制（储国正，1992），构造与岩浆活动的耦合共同制约了区内矿床的最终就位。

2.2 狮子山矿田

狮子山矿田位于铜陵矿集区东西向铜陵-南陵构造-岩浆带与北东向朱村复式向斜的交汇部位（图 3），是该矿集区内成矿规模最大的矿田（徐晓春等，2014）。矿田内出露地层发育完整，自志留系至第四系均有分布，其中石炭系至三叠系的碳酸盐岩地层是主要的赋矿围岩，为矽卡岩型矿床的形成提供了有利的岩性条件。岩浆活动以燕山期中酸性侵入岩为主。构造格架上，断裂系统较为发育，其中燕山期形成的北东向断裂与成矿关系最为密切（储国正，1992）。

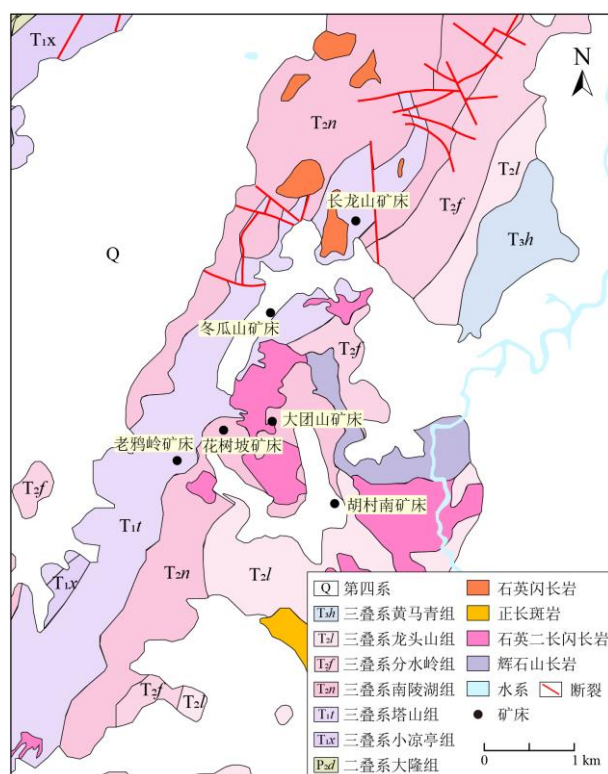


图 3 狮子山矿田地质图（据安徽省地质矿产勘查局 321 地质队修改）

Figure 3 Geological map of the Shizishan ore field (modified from the 321 Geological Party of Anhui Bureau of Geology and Mineral Resources).

狮子山矿田内矿床类型多样，浅部主要发育有浅成热液型及裂隙式和隐爆角砾岩筒式矽卡岩型矿床，中深部位主要发育接触式和层间式矽卡岩型矿床，深部发育层控式矽卡岩型矿床和斑岩型矿床，空间上呈现“三位一体”或“多层楼”的矿床空间分布模式（常印佛等，1991；翟裕生等，1992；徐晓春等，2014）。岩浆、构造、地层多个因素共同控制了狮子山矿田热液成矿系统，岩浆-流体系统充当了成矿物质的传输介质，褶皱带与断裂带则构成了

流体运移的主要输运通道,而有利的岩性地层往往成为成矿物质沉淀与富集的关键赋矿场所（黄许陈和储国正, 1993）。

2.3 典型矿床

矿田内发育狮子山、老鸦岭、花树坡、大团山及冬瓜山等多个热液型矿床。其中,冬瓜山矿床上部主要表现为层控式矽卡岩型矿床特征,矿体赋存在石炭系黄龙组和船山组地层内,深部则发育有斑岩型矿化;胡村矿床浅部矿体赋存于三叠系龙山组、南陵湖组与花岗闪长岩的接触部位,具有显著的接触交代型矿床特征,深部矿体发育在中二叠统栖霞组和孤峰组地层内,具有层控式矽卡岩型矿床特征。而东狮子山矿床主要发育在三叠系下统南陵湖组地层内及外接触带,地层被石英二长闪长岩和辉石闪长岩穿插、贯入和交代形成矽卡岩、大理岩和角岩,为隐爆角砾岩筒式矽卡岩型矿床（徐晓春等, 2014）。

3 三维成矿预测模型与预测信息集构建

3.1 三维地质建模

三维地质模型是开展深部隐伏矿体三维成矿预测的核心基础。本研究基于地质图、钻孔编录资料、物探反演剖面以及典型剖面地质信息等多源、多尺度地学信息,构建了狮子山矿田的三维地质模型。该模型清晰地揭示了区内成矿结构,显示了矿体赋存与石炭系—三叠系碳酸盐岩地层、中酸性岩体及断裂系统在空间上的密切耦合关系（图 4a）。此模型为后续的深部隐伏矿体预测与机器学习训练提供了统一的空间框架。

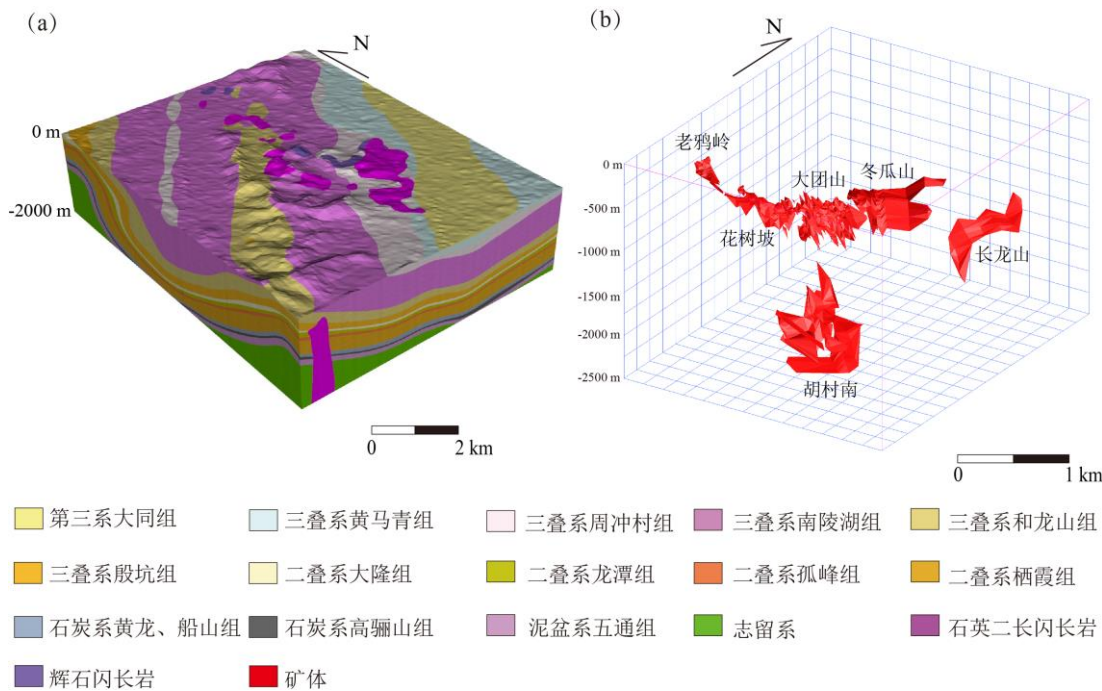


图 4 铜陵狮子山三维地质模型

(a) 狮子山矿田三维地质模型; (b) 典型矿床/矿段主要矿体模型

Figure 4 3D geological model of the Shizishan area, Tongling

(a) 3D geological model of the Shizishan ore field; (b) 3D model of main orebodies in typical deposits/ore blocks.

此外,研究还采用显式建模方法,对狮子山地区典型层控/接触交代型铜矿床的主要矿体的进行建模,定量分析三维成矿预测信息与矿化之间的空间关联。建模矿床包括老鸦岭铜

矿床、大团山铜矿床、花树坡铜矿床、冬瓜山铜矿床、胡村南铜矿床、长龙山铜矿床等。（图 4b）

3.2 三维成矿预测概念模型

在三维地质模型和已有数据支撑的基础上，研究建立了面向深部隐伏层控/接触交代型矽卡岩型矿床的三维成矿预测概念模型。模型构建遵循“地层—岩浆—构造”的预测模式，包含多种适用于三维成矿定量预测的三维预测要素（表 1）。

石炭系至三叠系的碳酸盐岩层位是矽卡岩交代作用的主要物质基础与矿质沉淀空间，控制了层控式矿体的产出；中酸性侵入岩体接触带是热液活动与化学交换的核心地带，直接控制了接触交代型矿体的定位；区域断裂系统为成矿流体提供了运移通道，而背斜核部等构造虚脱部位及 Si/Ca 地球化学界面则构成了流体汇聚与矿质富集的有利空间。

表 1 层控/接触交代型矽卡岩型矿床三维成矿预测概念模型

Table 1 Conceptual model of 3D mineral prospectivity mapping for stratabound and contact-metasomatic skarn deposits

类别	区域找矿预测标志	三维预测要素	三维空间分析方法
成矿有利地层	石炭系中上统黄龙-船山组、	各碳酸盐地层	三维地质体表面提取 三维距离场分析
	二叠系下统栖霞组、大隆组、		
	三叠系下统殷坑组、龙山组		
岩浆岩	闪长岩、石英二长/闪长斑岩、	中酸性岩体	三维地质体表面提取
	石英闪长岩等中酸性侵入岩	中酸性岩体表面	三维距离场分析
	背斜、向斜轴面	背斜、向斜轴面	
构造	北东-北北东向与北西向断裂	有利断裂	三维距离场分析
	Si/Ca 界面	Si/Ca 界面	

3.3 三维预测信息挖掘与预测数据集构建

为满足三维成矿预测需求，研究首先对三维地质模型进行网格化处理，网格单元设置为 100 m×100 m×50 m，形成包含 22.78 万个单元的规则化三维网格模型。该网格化模型完整保留了地层、构造和岩体的空间展布形态，以及相应的接触关系，为后续三维预测信息挖掘分析提供了高精度数据载体。

基于三维成矿预测概念模型和离散化后的三维地质模型，研究系统开展了三维空间分析，以挖掘可用于深部隐伏矿体预测的控矿信息，并构建预测数据集：研究采用三维地质体表面提取、三维距离场分析等三维空间分析技术，提取了碳酸盐岩地层界面、中酸性侵入岩体接触带、成矿相关断裂、Si/Ca 界面及褶皱轴面等关键地质界面（图 5）。为定量表征这些要素的空间控矿作用，本研究对所有地质界面进行了三维距离场计算，距离场定义了每个三维网格单元到最近目标地质界面的欧氏距离，从而将抽象的地质空间关系转化为连续的数值型变量；多个距离场的组合共同构成一个多维位置特征向量，这种连续、结构化的位置信息为后续全局注意力机制识别长距离地质关联提供了基础。

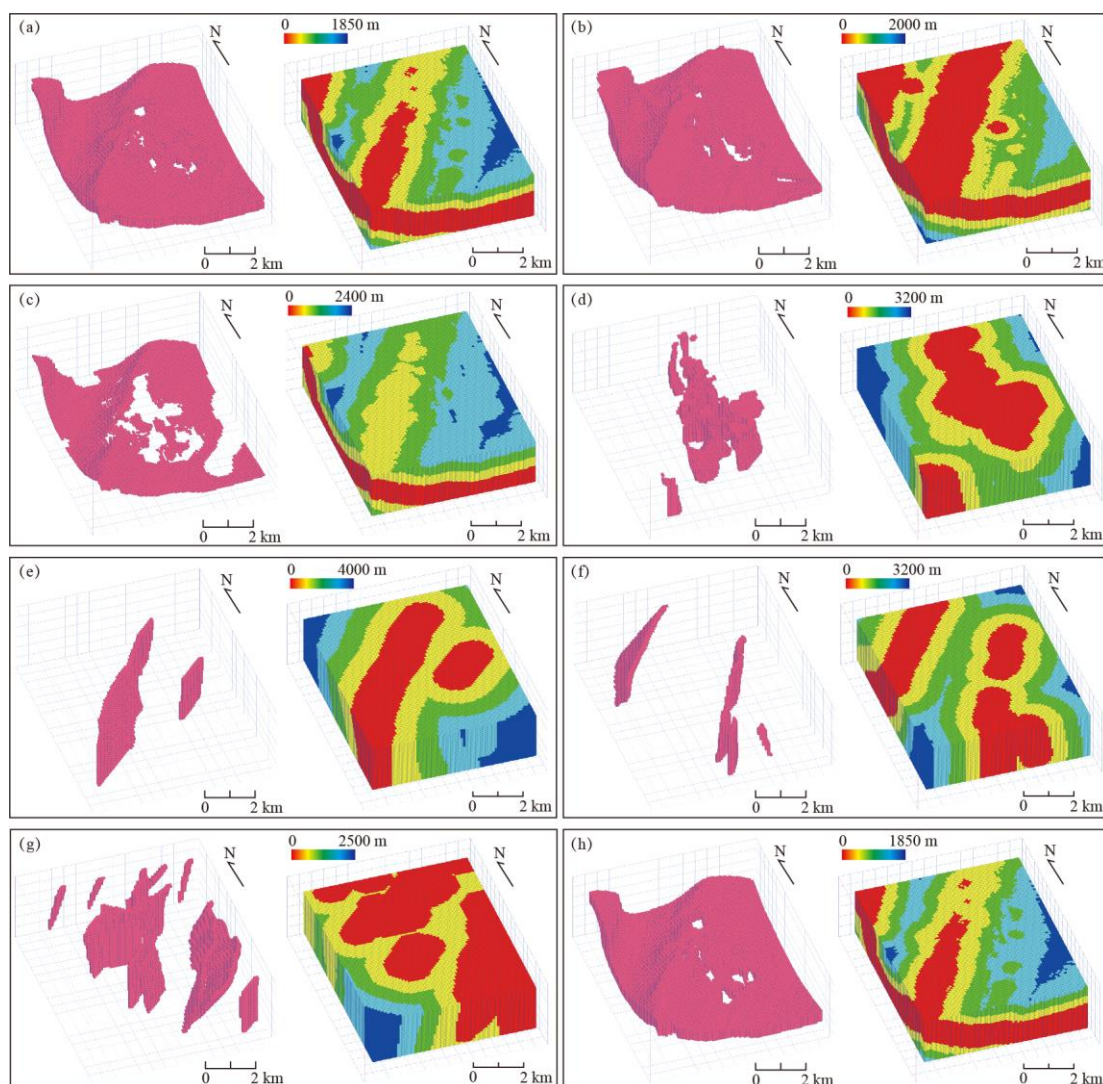


图 5 三维控矿要素

(a) 二叠系碳酸盐地层及三维距离场；(b) 三叠系碳酸盐地层及三维距离场；(c) 石炭系碳酸盐地层及三维距离场；(d) 成矿有利岩体及三维距离场；(e) 背斜轴面及三维距离场；(f) 向斜轴面及三维距离场；(g) 断裂及三维距离场；(h) Si/Ca 界面及三维距离场

Figure 5. Extraction of 3D ore-controlling factors

(a) Permian carbonate strata and 3D distance field; (b) Triassic carbonate strata and 3D distance field; (c) Carboniferous carbonate strata and 3D distance field; (d) Ore-favorable intrusions and 3D distance field; (e) Anticlinal axial surface and 3D distance field; (f) Synclinal axial surface and 3D distance field; (g) Faults and 3D distance field; (h) Si/Ca interface and 3D distance field

上述 8 个关键控矿要素作为特征，本研究进一步构建了结构化的三维成矿预测数据集。该数据集将作为后续机器学习模型的直接输入，为深部隐伏矿体的智能预测提供数据基础。

4 三维预测要素融合与靶区圈定

4.1 样本集构建

基于构建完成的预测数据集，进行正负样本集的构建。针对狮子山矿田层控/接触交代型矽卡岩矿床的成矿复杂性，研究以大团山、冬瓜山、花树坡、老鸦岭、长龙山和胡村南铜矿床内主要矿体作为数据驱动端元，将矿体内部网格单元定义为矿化单元，以每个矿体中心

点为核心，构建边长为 $9 \times 9 \times 9$ 网格单元的三维立方体样本块，用于表征矿化区的结构特征与空间组合关系。在剔除重复、边界不完整及信息缺失样本后，共获得 848 个正样本。为了保证负样本的地质背景与正样本具有可比性，本研究采用缓冲区分析法，在已知矿体外围一定范围内筛选无矿体显示的单元作为负样本来源，最终在研究区内构建 9653 个负样本。

为了确保模型的泛化能力，采用分层随机抽样策略，将正负样本以 8:2 的比例划分为训练集与测试集。最终，848 个正样本中 678 个被定义为训练集中的正样本，而剩余的则被指定为测试集中的正样本，9653 个负样本中随机选择了 678 个网格单元用于训练集，170 个负样本单元划入测试集。

4.2 模型训练

在训练过程中，研究采用二元交叉熵损失函数作为目标函数，并采用 Adam 优化器进行参数更新，初始学习率设定为 0.0001，训练过程共迭代 100 个轮次，并在每轮训练后使用测试集对模型性能进行验证，以监控模型收敛情况并防止过拟合。每轮训练结束后均在测试集上计算损失值与准确率，并以测试损失最小作为判定标准，自动保存性能最优的模型参数。在训练完成后加载该最优模型。

从准确率变化曲线（图 6）可以看出，在整个训练过程中，训练集与测试集的准确率逐渐增加。随着训练轮次的增加，准确率在训练后期趋于稳定，未观察到准确率显著下降的情况，表明模型在训练过程未发生过拟合现象。

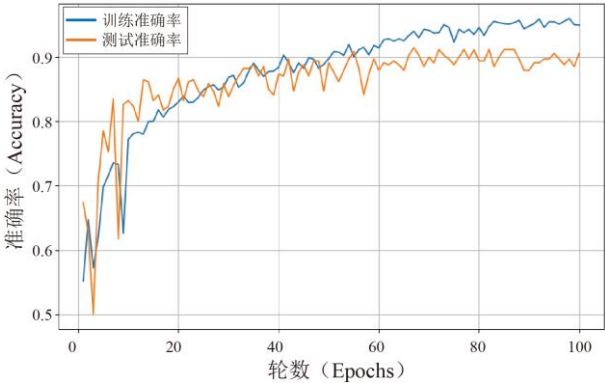


图 6 ViT 模型的训练和测试曲线

Figure 6 Training and testing curves of the ViT model

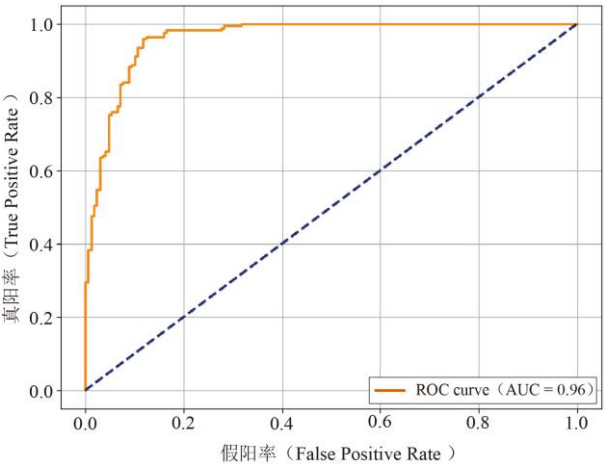


图 7 ViT 模型的 ROC 曲线

Figure 7 ROC curve of the ViT model

在完成模型训练并加载测试损失最优的参数后，本研究进行了系统的分类性能评价。结

果表明，3D-ViT 模型在三维控矿信息的识别方面具有较高的准确度与稳定性。从整体性能指标来看，模型的 ROC 曲线 AUC 值（Fawcett, 2006）为 0.9623，显示出良好的区分能力与预测可靠性，说明模型能够较为准确地区分矿化与非矿化单元（图 7）。

根据混淆矩阵结果，在 340 组测试样本中，模型共正确识别 161 个实际矿化样本，漏判 9 个矿化单元；同时正确识别 150 个非矿化样本，误将 20 个非矿化单元判为矿化。上述结果表明，在保持较高总体准确率的同时，模型对有矿类别具备更强的识别能力，能够有效响应隐伏矿体预测任务中对正样本识别性能的要求。基于混淆矩阵进一步计算的分类指标显示，模型的精确率为 0.8895、召回率为 0.9471、F1 分数为 0.9174。较高的召回率反映出模型漏判率较低，能够较为充分地捕捉深部隐伏矿体的空间位置；而良好的 F1 分数表明模型在正负样本不平衡背景下实现了精确率与召回率的有效权衡，保持了稳定的分类性能。

研究进一步采用捕捉效率曲线（Porwal et al., 2010），对模型的预测能力进行评价。结果显示（图 8），成矿概率最高的前 1.7% 的预测区域即可捕捉到 91% 的已知矿体；当考查前 3% 预测区域时，矿体捕捉率提升至 97.1%，表明模型能够有效学习已知矿化和三维预测信息之间的空间关联关系，展现出优异的预测能力。

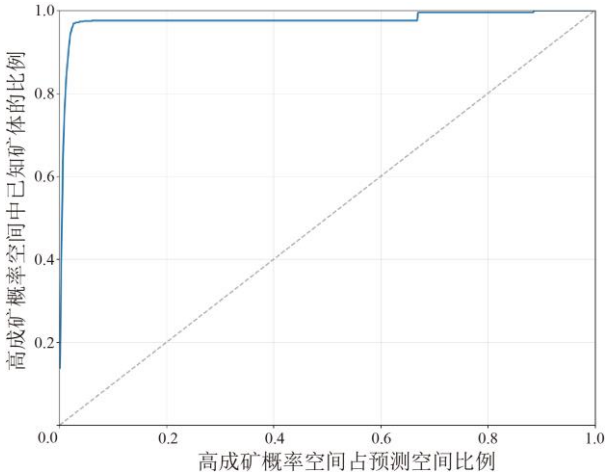


图 8 ViT 模型的捕获效率曲线

Figure 8 Capture efficiency curve of the ViT model

4.3 三维成矿预测结果

三维成矿概率分布图如图 9 所示。通过与实际矿体进行对比，模型不仅准确识别了典型矿体的主体位置，同时还进一步揭示出了层控矽卡岩矿体与接触交代矽卡岩矿体的主要分布特征。这说明 3D-ViT 能够有效捕捉不同控矿因素在三维空间中的关系，对复杂地质背景下的深部矿化信息具有较强的学习能力和表达能力。

在综合三维成矿概率分布、控矿因素空间约束及地质合理性分析的基础上，本研究最终在狮子山矿田深部圈定了 4 处找矿靶区。这些靶区的空间分布清晰地展示了该区层控-接触交代型矽卡岩成矿模式：层控型矿化主要受特定碳酸盐层位控制，矿体呈层状、似层状产出；接触交代型矿化则受中酸性岩体与围岩的接触带控制，矿体形态复杂多变。

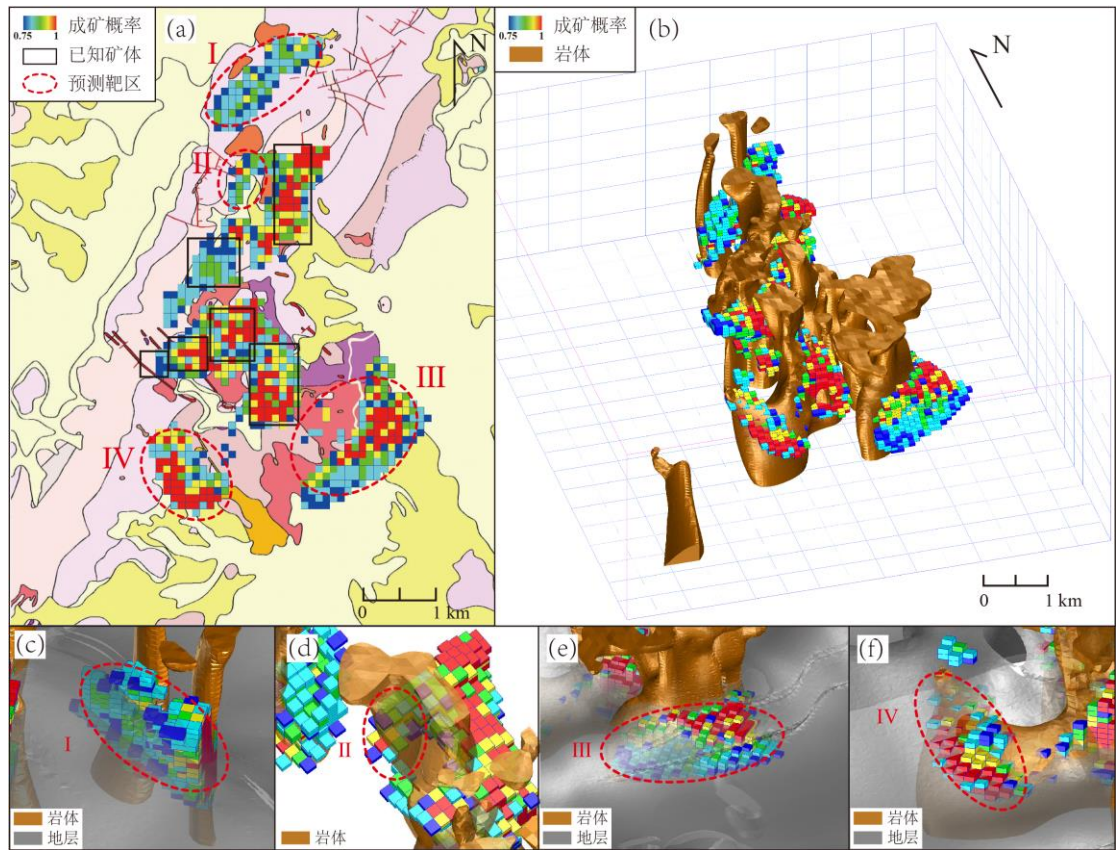


图 9 基于 ViT 模型的三维成矿预测结果图

(a) 俯视图; (b) 三维视图; (c) I号靶区; (d) II号靶区; (e) III号靶区; (f) IV号靶区

Figure 9 ViT-based 3D mineral prospectivity mapping results

(a) Plan view; (b) 3D perspective view; (c) Target Area I; (d) Target Area II; (e) Target Area III; (f) Target Area IV

I号与III号靶区具备层控特征，主要赋存于二叠系碳酸盐岩地层内，并紧邻中酸性岩体接触带（图 9c, e）。II号靶区：集中分布于岩体内部，表现为典型的接触交代型矽卡岩矿化区（图 9d）。IV号靶区：位于石炭系黄龙组与船山组碳酸盐岩内，一部分在岩体边部与地层接触区域，另一部分主要分布在深部隐伏岩体顶部，具备层控/接触交代型矽卡岩矿体的双重成矿特征（图 9f）。上述靶区的圈定充分体现了 ViT 模型在三维地质空间中识别矿化异常的能力，也展示出其在深部隐伏矿体预测中的应用潜力。

5 讨论

5.1 模型预测性能对比

为了验证 3D-ViT 模型在复杂三维控矿信息识别与深部隐伏矿体预测中，是否相较于传统机器学习方法具有预测能力方面的优势，本研究与多种传统机器学习模型（逻辑回归、随机森林、支持向量机）与深度学习模型（卷积神经网络）进行了对比分析。所有模型均在相同的数据集、相同训练测试划分及相同特征体系下进行训练，以确保结果的可比性与公平性。

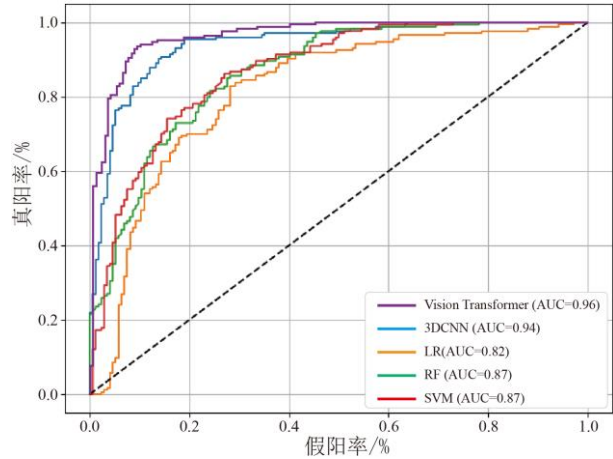


图 10 ViT、3DCNN、LR、RF 和 SVM 模型的 ROC 曲线

Figure 10 ROC curves of the ViT, 3D-CNN, LR, RF, and SVM models

综合表 2 与图 10 的评价指标，ViT 的预测性能显著优于传统机器学习模型。同时，与同为深度学习方法的 3D-CNN 相比，ViT 在召回率（Recall）与 F1 分数上展现了更优的均衡性与泛化能力。

表 2 各预测模型分类指标对比

Table 2 Comparison of classification metrics for different prediction models

模型	Accuracy	Precision	Recall	F1 Score	AUC
3D-ViT	0.9147	0.8895	0.9471	0.9174	0.9623
3D-CNN	0.8017	0.9412	0.6436	0.7645	0.9403
LR	0.7557	0.7170	0.8448	0.7757	0.8154
RF	0.7757	0.7566	0.821	0.7879	0.8650
SVM	0.7902	0.7512	0.8678	0.8053	0.8696

为评估模型在实际应用中的计算负担，所有实验均在同一配置平台（Intel Core i7-8700K CPU，NVIDIA GeForce GTX 1070 GPU，32GB 内存）上进行。结果显示，3D-ViT 模型训练耗时约 222 秒，预测整个研究区耗时约 84 秒；而作为对比的 3D-CNN 模型训练耗时约 359 秒，预测耗时约 45 秒。3D-ViT 在总耗时上低于 3D-CNN，表明通过合理的 3D Patch 划分与模型设计，其全局注意力机制并未导致计算开销的显著增加。

为直观对比不同模型的预测效果，图 11 展示了 3D-CNN 及传统机器学习模型在狮子山矿田的预测结果。对比可见，RF、LR 和 SVM 模型未能完整识别所有已知矿体，其预测的高概率区域存在明显遗漏，在捕捉深部隐伏矿化与复杂空间关联时存在固有局限。相比之下，3D-CNN 凭借其分层特征学习能力，成功检测到了全部已知矿体，但其高概率区的空间展布未能清晰表现层控型矿化的空间控制规律，缺乏与地层界面、岩体接触带等具体地质要素在形态和空间关系上的系统性耦合。

相比而言，ViT 模型借助其全局自注意力机制，能够同时建模地层、岩体与构造之间的远程空间关联，从而更清晰地刻画出层控型矿化对特定地层的依赖以及接触交代型矿化与岩体接触带的耦合规律，提升了预测的量化精度。

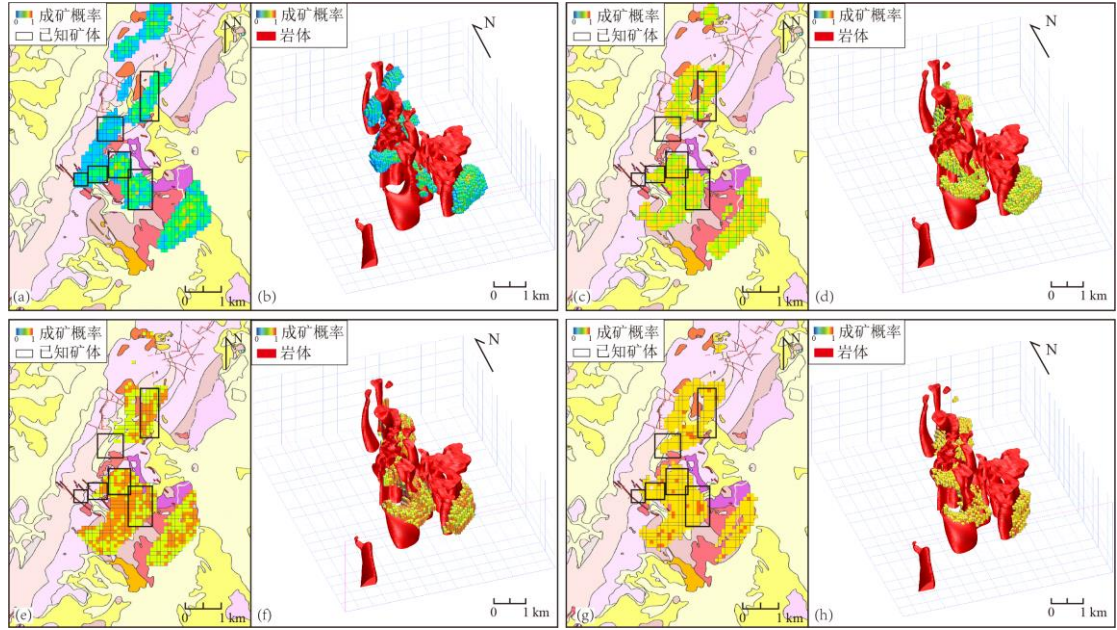


图 11 3DCNN (a)、LR (c)、RF (e) 和 SVM (g) 模型的预测结果俯视图；3DCNN (b)、LR (d)、RF (f) 和 SVM (h) 模型的预测结果三维视图

Figure 11 Prediction results of different models: Plan views of 3D-CNN (a), LR (c), RF (e), and SVM (g) models; 3D perspective views of 3D-CNN (b), LR (d), RF (f), and SVM (h) models.

5.2 模型结构分析

本研究提出的 ViT 模型,其核心优势在于能够系统性地建模多元控矿要素之间的远程依赖与全局关联,并在狮子山矿田的实例预测中得到了空间上的有效验证。

本研究加入的 3D Patch Embedding 模块与分离式三维位置编码,针对三维地质体数据,使模型能显式地感知并理解体素在深度、走向、倾向上的绝对位置与空间关系,为全局注意力计算提供了准确的空间基础。ViT 模型基于多头自注意力机制的全局建模能力,能够在超出单个卷积核感受野的空间尺度上,同时考量岩体接触带、特定地层与构造通道等多元控矿要素之间的长距离依赖关系。长距离依赖关系主要指在超出单个卷积核感受野的空间尺度上,不同地质要素之间存在的、控制矿体最终就位的成因或空间关联。这种关系无法通过局部特征的简单叠加来表征,必须通过全局性分析予以捕捉。

对于狮子山矿田,具有的多类型矿化在同一空间内不同位置的产出特征,深部隐伏的中酸性岩体的成矿流体,并非仅对其紧邻围岩发生作用,流体需要借助区域性的断裂通道或渗透性构造带,进行垂向与侧向运移,最终在远端特定碳酸盐岩层位中沉淀成矿。本研究所构建的 ViT 模型在评估任一矿体成矿潜力时,同步考量全区范围内的特征信息。预测结果印证了模型对此类关系的捕捉能力,靶区不仅出现在岩体接触带附近,也在远离岩体但符合特定地层的深部区域呈现高概率。这表明,3D-ViT 不仅学习了局部的找矿标志,更通过其全局注意力机制,识别并建模了这些控制矿床空间定位的长距离地质依赖关系。

尽管 ViT 模型具有较高的准确性,但其整体决策过程依然不透明,模型最终预测结果难以解释,这是当前基于注意力机制的深度学习模型在应用于解释性任务时面临的共同挑战 (Jain and Wallace, 2019)。本研究的主要贡献在于验证了 3D-ViT 框架在捕捉复杂三维空间依赖与提升预测精度方面的有效性与可行性。如何深入剖析注意力权重的地质语义,发展针对地质 Transformer 模型的可解释性方法,将是连接数据驱动预测与地质成因机理理解的关键研究方向。

结论

- (1) 本研究提出了能够有效应用于深部隐伏矿床三维成矿预测的 ViT 模型架构。实例结果显示,相比多种浅层机器学习模型和深度学习三维卷积神经网络模型,本文构建的 ViT 模型在三维信息表达、远程空间关联提取方面表现更优,并在多项分类指标上均取得最佳结果,显示其在复杂地质条件下开展隐伏矿体三维预测的有效性,并具有良好的研究和应用潜力。
- (2) 基于构建的 ViT 模型,研究在铜陵狮子山矿田开展了三维成矿预测,圈定 4 处找矿靶区;模型预测结果具有较强的地质合理性,为矿田矽卡岩型矿床的进一步找矿工作提供了方向和依据。

References

- Chang, Y.F., Liu, X.P., Wu, Y.C., 1991. The Copper-Iron Belt of the Lower and Middle Reaches of the Changjiang River. Geological Publishing House, Beijing (in Chinese with English abstract).
- Chu, G.Z., 1992. Shizishan Orefield Tectonics and the Characteristics of Its Control Over Rocks and Ores. *Geology of Anhui*, 2(2): 1-14 (in Chinese with English abstract).
- Deng, H., Zheng, Y., Chen, J., 2022. Learning 3D Mineral Prospectivity from 3D Geological Models Using Convolutional Neural Networks: Application to a Structure-Controlled Hydrothermal Gold Deposit. *Computers & Geosciences*, 161: 105074. doi: 10.1016/j.cageo.2022.105074
- Dosovitskiy, A., 2020. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv preprint arXiv:2010.11929*. doi: 10.48550/arXiv.2010.11929
- Fawcett, T., 2006. An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27(8): 861-874. doi: 10.1016/j.patrec.2005.10.010
- Fu, Z.J., Li, J.J., Ren, L., et al., 2023. SLDDNet: Staged Short and Long Distance Dependency Network for Remote Sensing Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-19. doi: 10.1109/TGRS.2023.3305554
- Gao, L., Gopalakrishnan, G., Nasri, A., et al., 2025. Transformer-GCN Fusion Framework for Mineral Prospectivity Mapping: A Geospatial Deep Learning Approach. *Minerals*, 15(7): 711. doi: 10.3390/min15070711
- Huang, X.C., Chu, G.Z., 1993. Multistory Metallogenic Model of the Shizishan Orefield in Tongling, Anhui Province. *Mineral Deposits*, 12(3): 221-230 (in Chinese with English abstract).
- Jain, S., Wallace, B. C., 2019. Attention Is Not Explanation. *arXiv preprint arXiv:1902.10186*. doi: 10.48550/arXiv.1902.10186
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 60(6): 84-90. doi: 10.1145/3065386
- Li, C., Xiao, K.Y., Sun, L., et al., 2024. CNN-Transformers for Mineral Prospectivity Mapping in the Maodeng-Baiyinchagan Area, Southern Great Xing'an Range. *Ore Geology Reviews*, 167: 106007. doi: 10.1016/j.oregeorev.2024.106007
- Li, H., Li, X.H., Yuan, F., et al., 2020. Convolutional Neural Network and Transfer Learning

- Based Mineral Prospectivity Modeling for Geochemical Exploration of Au Mineralization within the Guandian–Zhangbaling Area, Anhui Province, China. *Applied Geochemistry*, 122: 104747. doi: 10.1016/j.apgeochem.2020.104747
- Li, X.H., Chen, Y.H., Yuan, F., et al., 2024. 3D Mineral Prospectivity Modeling Using Multi-Scale 3D Convolution Neural Network and Spatial Attention Approaches. *Geochemistry*, 84(4): 126125. doi: 10.1016/j.chemer.2024.126125
- Li, X.H., Xue, C., Chen, Y.H., et al., 2023. 3D Convolutional Neural Network-Based 3D Mineral Prospectivity Modeling for Targeting Concealed Mineralization within Chating Area, Middle-Lower Yangtze River Metallogenic Belt, China. *Ore Geology Reviews*, 157: 105444. doi: 10.1016/j.oregeorev.2023.105444
- Li, X.H., Yuan, F., Zhang, M.M., et al., 2015. Three-Dimensional Mineral Prospectivity Modeling for Targeting of Concealed Mineralization within the Zhonggu Iron Orefield, Ningwu Basin, China. *Ore Geology Reviews*, 71: 633-654. doi: 10.1016/j.oregeorev.2015.06.001
- Mao, J.W., Shao, Y.J., Xie, G.Q., et al., 2009. Mineral Deposit Model for Porphyry-Skarn Polymetallic Copper Deposits in Tongling Ore Dense District of Middle-Lower Yangtze Valley Metallogenic Belt. *Mineral Deposits*, 28(2): 109-119 (in Chinese with English abstract).
- Mao, X.C., Su, Z., Deng, H., et al., 2024. Three-Dimensional Mineral Prospectivity Modeling with Geometric Restoration: Application to the Jinchuan Ni – Cu – (PGE) Sulfide Deposit, Northwestern China. *Natural Resources Research*, 33(1): 75-105. doi: 10.1007/s11053-023-10269-2
- Mao, X.C., Zhang, M.M., Deng, H., et al., 2016. Three-Dimensional Visualization Prediction Method for Concealed Ore Bodies in Deep Mining Areas. *Journal of Geology*, 40(3): 363-371 (in Chinese with English abstract).
- Ning, Y., Wang, Y.Z., Lu, J.L., et al., 2025. Mineral Prospectivity Mapping for Multi-Source Geoscience Data: A Novel Unsupervised Deep Learning Method. *Ore Geology Reviews*, 186: 106866. doi: 10.1016/j.oregeorev.2025.106866
- Porwal, A., Gonzalez-Alvarez, I., Markwitz, V., et al., 2010. Weights-of-Evidence and Logistic Regression Modeling of Magmatic Nickel Sulfide Prospectivity in the Yilgarn Craton, Western Australia. *Ore Geology Reviews*, 38(3): 184-196. doi: 10.1016/j.oregeorev.2010.04.002
- Singer, D.A., 2018. Comparison of Expert Estimates of Number of Undiscovered Mineral Deposits with Mineral Deposit Densities. *Ore Geology Reviews*, 99: 235-243. doi: 10.1016/j.oregeorev.2018.06.019
- Sun, T., Li, H., Wu, K.X., et al., 2020. Data-Driven Predictive Modelling of Mineral Prospectivity Using Machine Learning and Deep Learning Methods: A Case Study from Southern Jiangxi Province, China. *Minerals*, 10(2): 102. doi: 10.3390/min10020102
- Vaswani, A., Shazeer, N., Parmar, N., et al., 2017. Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30. doi: 10.48550/arXiv.1706.03762
- Wang, G.W., Li, R.X., Carranza, E.J.M., et al., 2015. 3D Geological Modeling for Prediction of Subsurface Mo Targets in the Luanchuan District, China. *Ore Geology Reviews*, 71: 592-610. doi: 10.1016/j.oregeorev.2015.03.002
- Wang, J.J., Xie, H.R., Wang, F.L., et al., 2023. A Transformer–Convolution Model for Enhanced Session-Based Recommendation. *Neurocomputing*, 531: 21-33. doi:

10.1016/j.neucom.2023.01.083

- Xiao, F., Cheng, Q.M., Hou, W.S., et al., 2025. Three-Dimensional Prospectivity Modeling of Jinshan Ag-Au Deposit, Southern China by Weights-of-Evidence. *Journal of Earth Science*, 36(5): 2038 – 2057. doi: 10.1007/s12583-023-1822-6
- Xu, X.C., Fan, Z. L., He, J., et al., 2014. Metallogenic Model for the Copper-Gold-Polymetallic Deposits in Shizishan Ore-Field, Tongling, Anhui Province. *Acta Petrologica Sinica*, 30(4): 1054-1074 (in Chinese with English abstract).
- Xu, Y., Zuo, R.G., Chen, Z.Y., et al., 2025. Recent Advances and Future Research Directions in Deep Learning as Applied to Geochemical Mapping. *Earth-Science Reviews*, 105209. doi: 10.1016/j.earscirev.2025.105209
- Xu, Y.Y., Li, Z.X., Xie, Z., et al., 2021. Mineral Prospectivity Mapping by Deep Learning Method in Yawan-Daqiao Area, Gansu. *Ore Geology Reviews*, 138: 104316. doi: 10.1016/j.oregeorev.2021.104316
- Yang, F.F., Zuo, R.G., 2024. Geologically Constrained Convolutional Neural Network for Mineral Prospectivity Mapping. *Mathematical Geosciences*, 56(8): 1605-1628. doi: 10.1007/s11004-024-10141-w
- Yousefi, M., Lindsay, M.D., Kreuzer, O., 2024. Mitigating Uncertainties in Mineral Exploration Targeting: Majority Voting and Confidence Index Approaches in the Context of an Exploration Information System (EIS). *Ore Geology Reviews*, 165: 105930. doi: 10.1016/j.oregeorev.2024.105930
- Yuan, F., Li, X.H., Zhang, M.M., et al., 2018. Research Progress of 3D Prospectivity Modeling. *Gansu Geology*, 27(1): 32-36 (in Chinese with English abstract).
- Zhai, M.G., Wu, F.Y., Hu, R.Z., et al., 2019. Critical Metal Mineral Resources: Current Research Status and Scientific Issues. *Bulletin of National Natural Science Foundation of China*, 33(2): 106-111 (in Chinese with English abstract).
- Zhai, Y.S., Yao, S.Z., Lin, X.D., et al., 1992. Metallogenic Regularity of Iron and Copper Deposits in the Middle and Lower Valley of the Yangtze River. *Mineral Deposits*, 11(1): 1-12 (in Chinese with English abstract).
- Zhang, S., Carranza, E.J.M., Wei, H.T., et al., 2021. Data-Driven Mineral Prospectivity Mapping by Joint Application of Unsupervised Convolutional Auto-Encoder Network and Supervised Convolutional Neural Network. *Natural Resources Research*, 30(2): 1011-1031. doi: 10.1007/s11053-020-09789-y
- Zhou, T.F., Fan, Y., Wang, S.W., et al., 2017. Metallogenic Regularity and Metallogenic Model of the Middle-Lower Yangtze River Valley Metallogenic Belt. *Acta Petrologica Sinica*, 33(11): 3353-3372 (in Chinese with English abstract).
- Zuo, R.G., 2025. Key Technology for Intelligent Mineral Prospectivity Mapping: Challenges and Solutions. *Science China Earth Sciences*, 68(9): 2976-2991. doi: 10.1007/s11430-025-1622-1
- Zuo, R.G., Yang, F.F., Cheng, Q.M., et al., 2025. A Novel Data-Knowledge Dual-Driven Model Coupling Artificial Intelligence with a Mineral Systems Approach for Mineral Prospectivity Mapping. *Geology*, 53(3): 284-288. doi: 10.1130/G52970.1

附中文参考文献

常印佛, 刘湘培, 吴言昌, 1991.长江中下游铜铁成矿带.北京: 地质出版社.

- 储国正, 1992.铜陵狮子山矿田构造及其控岩控矿作用的研究.安徽地质, (2):1-14.
- 黄许陈, 储国正, 1993.铜陵狮子山矿田多位一体(多层楼)模式.矿床地质, (03):221-230.
- 毛景文, 邵拥军, 谢桂青, 等, 2009.长江中下游成矿带铜陵矿集区铜多金属矿床模型.矿床地质, 28(02):109-119.
- 毛先成, 张苗苗, 邓浩, 等, 2016.矿区深部隐伏矿体三维可视化预测方法.地质学刊, 40(03):363-371.
- 徐晓春, 范子良, 何俊, 等, 2014.安徽铜陵狮子山矿田铜金多金属矿床的成矿模式.岩石学报, (4):1054-1074.
- 袁峰, 李晓晖, 张明明, 等, 2018.三维成矿预测研究进展.甘肃地质, 27(1):32-36.
- 翟明国, 吴福元, 胡瑞忠, 等, 2019.战略性关键金属矿产资源:现状与问题.中国科学基金, 33(2):106-111.
- 翟裕生, 姚书振, 林新多, 等, 1992.长江中下游地区铁、铜等成矿规律研究.矿床地质, (01):1-12.
- 周涛发, 范裕, 王世伟, 等, 2017.长江中下游成矿带成矿规律和成矿模式.岩石学报, 33 (11): 3353-3372.