

## 融合形状变换及纹理学习的肺结节生长预测

马力, 黄德皇, 王艳芳

### Predicting Lung Nodule Growth with Shape Transformation and Texture Learning

MA Li, HUANG Dehuang, and WANG Yanfang

在线阅读 View online: <https://doi.org/10.15953/j.ctta.2023.167>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### CT纹理特征分析在孤立性肺结节诊断中的研究进展

Advance of CT Texture Feature Analysis in Diagnosis of Solitary Pulmonary Nodules

CT理论与应用研究. 2020, 29(1): 111–118

#### 初探肺纯磨玻璃结节3年随访期内增长的风险因素

Explore the Risk Factors for the Growth of Pure Lung Ground Glass Nodules During the 3-year Follow Up Period

CT理论与应用研究. 2019, 28(5): 617–624

#### 基于CT表现的孤立性肺结节良恶性预测模型的研究

A Predicting Model to Estimate the Probability of Malignancy in Solitary Pulmonary Nodules Basing on CT Images

CT理论与应用研究. 2019, 28(6): 677–683

#### I期浸润性肺腺癌磨玻璃影及附壁样生长三维CT值定量分析

3D CT Value Analysis of Ground Glass Shadow and Coelentereous Growth in Stage I Invasive Lung Adenocarcinoma

CT理论与应用研究. 2018, 27(1): 115–121

#### 基于人工智能的胸部CT肺结节检出及良恶性诊断效能评估

Evaluation of Detection and Diagnostic Efficiency of Pulmonary Nodules by Chest CT Based on Artificial Intelligence

CT理论与应用研究. 2021, 30(6): 709–715

#### 探讨CT扫描剂量对人工智能检测肺结节效能的影响

To Explore the Effect of CT Scan Dose on the Efficacy of Artificial Intelligence in Detecting Lung Nodules

CT理论与应用研究. 2021, 30(4): 455–465



关注微信公众号, 获得更多资讯信息

马力, 黄德皇, 王艳芳. 融合形状变换及纹理学习的肺结节生长预测 [J]. CT理论与应用研究 (中英文), 2024, 33(3): 317-324. DOI:[10.15953/j.ctta.2023.167](https://doi.org/10.15953/j.ctta.2023.167).

MA L, HUANG D H, WANG Y F. Predicting Lung Nodule Growth with Shape Transformation and Texture Learning [J]. CT Theory and Applications, 2024, 33(3): 317-324. DOI:[10.15953/j.ctta.2023.167](https://doi.org/10.15953/j.ctta.2023.167). (in Chinese).

# 融合形状变换及纹理学习的肺结节生长预测

马力<sup>1</sup>, 黄德皇<sup>2</sup>, 王艳芳<sup>1✉</sup>

1. 中山仰视科技有限公司, 广东 中山 528400

2. 中山北京理工大学研究院, 广东 中山 528400

**摘要:** 虽然人工智能在肺结节检测方面已经相当成熟, 但对其生长预测的研究仍然有限。准确的生长预测有助于临床决策, 为患者随访策略提供信息。本文提出一种新的结节生长预测网络模型, 该模型可以在特定时间间隔生成高质量的肺结节图像。模型使用双分支结构对肺结节图像进行特征提取, 其中一个分支, 利用位移场预测机制, 通过体素水平的未来位移估计来学习肺结节的形状转换; 另一分支, 采用3DU-Net, 学习肺结节的纹理变化。随后, 对提取的高维特征图通过坐标注意力机制, 突出有利的图像特征, 再拼接两个分支的结果, 输入至特征重建模块得到最终的肺结节生长预测图像。同时, 本文引入时间间隔编码模块, 将期望的时间间隔纳入网络, 从而能够生成不同未来时间点的预测图像。

**关键词:** 肺结节; 生长预测; 位移场; 时间间隔编码

DOI:[10.15953/j.ctta.2023.167](https://doi.org/10.15953/j.ctta.2023.167) 中图分类号: O 242; R 814; R 563 文献标识码: A

肺癌是所有癌症中最常被诊断出的癌症。据世界卫生组织统计, 2020年肺癌造成约180万人死亡。据估计, 这一数字在未来的几年中仍会持续不断地增加。此外, 肺癌患者的五年生存率大约19.7%<sup>[1]</sup>, 肺癌是由肺结节生长发展而形成, 如果高危肺结节能够早期被发现并及时随访治疗, 往往是可以成功控制病情甚至治愈。但在大多数情况下, 当患者最终被诊断出肺癌时, 病情早已进入晚期阶段, 肺部恶性结节的生长已经不可逆转, 即使是经验丰富的医生也没有太好的办法, 在这种情况下治疗通常都收效甚微并会极大增加患者的痛苦。通过对高危人群进行定期的影像学筛查, 对可疑恶性的肺结节进行较为密切的随访, 再根据其在随访期间的变化给予临床指导是非常可取的。

为此, 各种计算机辅助诊断系统<sup>[2-3]</sup>被开发出来, 以帮助医生更好的检出肺结节并对其良恶性分析, 而对结节的生长预测功能的需求随着检测系统的成熟变得更为迫切。

## 1 相关工作

基于深度学习的肺结节生长预测研究相对检测和分割任务而言较少, 对肺结节生长预测研究, 可用的训练标注数据缺乏是影响其研究进展的重要因素。

Zhang等<sup>[4]</sup>利用深度卷积神经网络直接表示和学习细胞侵袭和质量效应的可能性, 并预测肿瘤的后续可能关联区域, 该方法输入需要结合3种多模态数据, 模型输入限制较多, 且需计算光流, 网络计算量大。Rafael-Palou等<sup>[5]</sup>提出了一种3D孪生神经网络模型方法, 用于在同一患者的两次CT扫描中重新识别结节, 而无需图像配准, 该网络被集成到一个两阶段的自动管道中, 以检测、匹配和预测给定对CT扫描的结节生长, 独立测试集的结果显示, 结节检测灵敏度为94.7%, 该方法只能对结节是否生长进行预测, 而无法生成高质量的结节生长预测图像来直观的观测结节的生长图像。Sheng等<sup>[6]</sup>中提出了一种统一的肺结节随访视觉预测与诊断框架, 通过空间变换, 利用卷积网络对连续CT扫描之间的结节生长进行建模, 预测未来结节的生长, 可根据预测的结节生长情况和以前的诊

收稿日期: 2023-08-29。

基金项目: 中山市2019年高端科研机构创新专项(第一批)(基于人工智能CT时序列的肺癌早期预测及其应用)。

断对进行结节分类, 该方法收集成对的训练数据时, 并未考虑时间间隔作为结节生成的控制条件, 无法对生成的结节进行时间维度的控制。Li 等<sup>[7]</sup>提出了 NoFoNet 网络, 该模型通过利用肺结节分割掩码及肺结节图像作为输入, 学习结节掩码形状变换及结节的纹理变化, 基于分割结果限定结节的生成区域, 生成了高质量的肺结节预测图像, 但该方法输入部分需要额外引入结节分割掩码。

对上述方法存在的问题, 本文提出一种结节生长预测网络模型, 无需结节分割掩码, 即可预测生成任意指定时间间隔的肺结节预测图像, 且生成的结果具有较高的质量。

本文主要贡献: ① 使用双分支结构学习形状变换及纹理特征, 其中一个分支通过预测每个体素的未来位移场, 作用于输入图像来学习肺结节形状变换, 另一分支, 采用 3D U-Net<sup>[8]</sup>学习肺结节的纹理变化; ② 采用坐标注意力机制分别作用于提取的双分支特征图, 用于突出有利的图像特征并进行拼接融合; ③ 引入特征重建模块, 对融合的特征进行重构, 生成最终的预测结果; ④ 引入时间间隔编码模块, 对时间间隔进行编码加入至网络模型, 用以从时间维度控制生成不同时间间隔的预测图像。

## 1.1 网络整体结构

为了模拟结节的生长, 本文提出一个结节生长预测网络 (nodule growth prediction net, NGPNet)。该网络由形状及纹理特征提取模块 (shape and texture feature extraction module, STFEM), 坐标注意力模块 (coordinate attention module, CAM), 特征重建模块 (feature reconstruction module, FRM) 及时间编码模块 (time encoding module, TEM), 4 个核心模块组成。NGPNet 整体的网络结构 (图 1)。

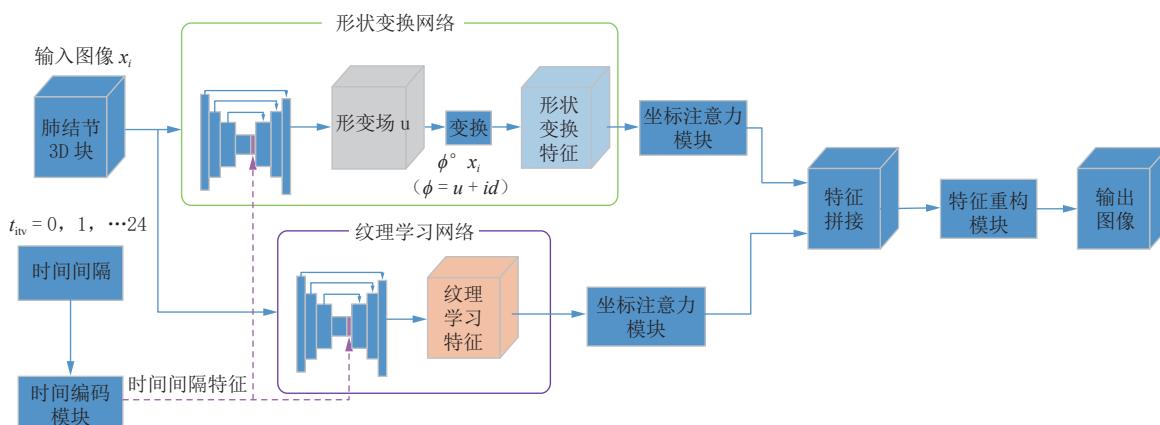


图 1 网络模型整体结构

Fig. 1 The overall structure of the network model

首先, 将肺结节体素块输入到形状变换分支和纹理学习分支, 进行形状变换和纹理特征的提取, 对提取的特征图通过坐标注意力机制, 突出有利的图像特征。然后拼接两个分支的结果, 输入至特征重建模块得到最终的肺结节生长预测图像。为了生成不同随访间隔的图像, 模型将时间间隔信息编码嵌入至网络模型。

模型训练时需要给定一对肺结节的随访图像作为网络的输入图像及目标图像, 同时需要提供该对结节随访时间间隔, 通过时间编码模块将时间间隔编码成 embedding 特征加入网络, 进行时间条件的学习, 用以推理时能从时间维度控制生成不同时间间隔的结果。

## 1.2 时间编码模块 (TEM)

由于两次随访 CT 检查的时间可能不同, 需要对随访时间间隔作为约束条件引入网络模型, 对生成的图像进行时间上的控制。根据 Li 等<sup>[7]</sup>的方法, 进行时间间隔编码。由于本文研究的数据集大小的限制, 且大多数时间间隔小于 24 个月, 时间间隔取值为 0 到 24 的离散整数数值, 表示随访间隔月份, 超过 24 个月用 24 代替。在 TEM 中使用不同频率的正弦和余弦函数来编码时间信息:

$$F_{\text{TE}}(t_{\text{itv}}, 2i) = \sin\left(t_{\text{itv}}/100^{\frac{2i}{d_{\text{fm}}}}\right), \quad (1)$$

$$F_{\text{TE}}(t_{\text{itv}}, 2i+1) = \cos\left(t_{\text{itv}}/100^{\frac{2i}{d_{\text{fm}}}}\right), \quad (2)$$

式中  $t_{\text{itv}}$  为时间间隔,  $d_{\text{fm}}$  为编码特征映射的总通道数,  $i$  为特征维数。时间编码的奇偶维是由不同波长 ( $2\pi$  到  $100 \times 2\pi$ ) 的  $\sin/\cos$  函数生成的, 这使得相对时间信息以冗余的方式编码。此外, 由于  $\sin/\cos$  函数的有界性, 编码结果的取值范围在一定的数值区间内。这两点保证了时间编码方法能够生成更有意义的高维表示空间。

### 1.3 形状及纹理特征提取模块 (STFEM)

对形状变换网络分支, 该方法的核心参考论文<sup>[7]</sup>, 通过预测一个位移场来进行后续形状变换。该方法借鉴配准模型<sup>[9]</sup>的思路, 分支首先使用 3D U-Net 预测一个平滑的体素方向位移场  $u$ :

$$u = F_{w1}\left(\left(x_i, F_{\text{TE}}(t_{\text{itv}})\right); \theta_{w1}\right), \quad (3)$$

其中  $F_{w1}$  代表参数为  $w1$  的 3D U-Net, 用来预测位移场,  $x_i$  为输入图像,  $F_{\text{TE}}(t_{\text{itv}})$  表示时间间隔  $t_{\text{itv}}$  通过时间编码模块得到的 embedding 特征, 该特征被嵌入至 3D U-Net 结构的底部 (图 1)。

然后进行空间形状变换。按照配准模型<sup>[9]</sup>, 有变换函数  $\varphi = u + id$ , 其中  $id$  为恒等函数。将变换函数  $\varphi$  作用于输入图像  $x_i$  得到形状变换图像  $x_w$ , 可表示为:

$$x_w = \Phi \odot x_i. \quad (4)$$

对纹理学习分支, 使用 3D U-Net 直接学习纹理细节, 学习的特征  $f$  可表示为:

$$f = F_{w2}\left(\left(x_i, F_{\text{TE}}(t_{\text{itv}})\right); \theta_{w2}\right), \quad (5)$$

其中  $F_{w2}$  代表参数为  $w2$  的 3D U-Net, 用来学习纹理细节,  $x_i$  为输入图像,  $F_{\text{TE}}(t_{\text{itv}})$  表示时间间隔  $t$  通过时间编码模块得到的 embedding 特征, 该特征被嵌入至 3D U-Net 结构的底部。

### 1.4 坐标注意力模块 (CA)

在计算机视觉任务中, 注意力机制可以帮助模型更好地理解输入图像或视频, 从而提高模型的性能。常用的通道 SE (squeeze-and-excitation, SE) 注意力<sup>[10]</sup>会忽略位置信息, 而位置信息对于生成空间选择性注意力图非常重要。本文参考坐标注意力 (coordinate attention, CA)<sup>[11]</sup> 模块通过精确的位置信息对通道关系和长程依赖进行编码, 使网络能够以较小的计算成本关注大的重要区域。坐标注意力将通道注意力分解为两个一维特征编码, 分别沿两个空间方向聚合特征。这样既可以获取一个空间方向上的远程依赖关系, 又可以保留另一个空间方向上的精确位置信息。然后, 生成的特征图被分别编码成一对方向感知和位置敏感的注意力图, 这对注意力图可以互补地应用于输入特征图, 以增加感兴趣对象的表示。CA 模块细节如图 2 所示, 具体分成坐标信息嵌入和坐标注意力生成两部分。

**坐标信息嵌入:** 对提取到的形状变换或纹理特征  $X$ ,  $X \in \mathbb{R}^{N \times X \times C \times H \times W}$ , 先使用尺寸  $(H, 1)$  和  $(1, W)$  的池化核沿着水平坐标方向和竖直坐标方向对每个通道进行编码。对第  $C$  个通道, 在  $H$  方向上的输出可用公式 (6) 表示,

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i). \quad (6)$$

同理, 对第  $C$  个通道, 在  $W$  方向上的输出可用公式 (7) 表示,

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w). \quad (7)$$

上面两个变换沿着两个空间方向进行特征聚合，得到一对方向感知注意力图。这两种变换允许注意力模块捕捉沿着一个空间方向的长期依赖关系，并保存沿着另一个空间方向的精确位置信息，有助于模型更好地定位感兴趣目标。

**坐标注意力生成：**对公式(6)和公式(7)生成的聚合特征图，可以很好的获得全局感受野对于精确位置信息的编码。为了利用产生的特征，通过信息嵌入中的变换后，对公式(6)和公式(7)生成的聚合特征图进行拼接，然后利用一个共享的  $1 \times 1$  卷积变换函数  $F_1$ ，得到

$$f = \delta(F_1([z^h, z^w])), \quad (8)$$

其中  $[\cdot, \cdot]$  表示沿空间维度的拼接操作， $\delta$  为非线性激活函数， $f \in \mathbb{R}^{N \times T \times \frac{C}{r} \times (H+W)}$  为从水平方向和垂直方向编码空间信息得到的中间特征图。然后沿着空间维度把  $f$  分解成两个单独的张量  $f^h \in \mathbb{R}^{N \times T \times \frac{C}{r} \times H}$  和  $f^w \in \mathbb{R}^{N \times T \times \frac{C}{r} \times W}$ ，再利用两个  $1 \times 1$  卷积变换函数  $F_h$  和  $F_w$  分别将张量变换回输入  $X$  相同的通道数，得到  $g^h, g^w$ 。

$$g^h = \delta(F_h(f^h)), \quad (9)$$

$$g^w = \delta(F_w(f^w)). \quad (10)$$

输出  $g^h, g^w$  作为注意权重，最终的空间注意力模块的输出为：

$$F_o = X + X \times g^h \times g^w, \quad (11)$$

其中  $F_o \in \mathbb{R}^{N \times T \times C \times H \times W}$ 。

对双分支网络输出特征图分别使用坐标注意力模块进行感兴趣区域特征提取，将所得的特征图进行拼接操作，输入至特征重建模块，进行进一步的细节重建。

## 1.5 特征重建模块

特征重建模块由 2 个基础的 3D Residue Block<sup>[12]</sup> 残差单元及 2 个 Conv-BN-ReLU 结构组成，用以精细化的调整重建的生成细节。

## 1.6 损失函数

本文损失函数采用基于 SSIM<sup>[13]</sup> 的结构损失和基于 L2 范数的内容损失的多损失函数。结构相似性度量指标 SSIM 从亮度、对比度和结构 3 方面衡量目标图像与预测图像相似性程度。设  $x$  为目标图像， $y$  为预测图像，其数学表达式为：

$$\text{SSIM}_{x,y} = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \cdot \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \quad (12)$$

式中， $\mu$  和  $\sigma$  分别表示均值和标准差， $\sigma_{xy}$  表示  $x, y$  的协方差， $C_1, C_2$  和  $C_3$  为常数系数。以结构相似性 SSIM 为基础，针对预测图像的结构损失  $L_{\text{SSIM}}$ ， $G$  和  $P$  分别表示目标图像和预测图像，则  $L_{\text{SSIM}}$

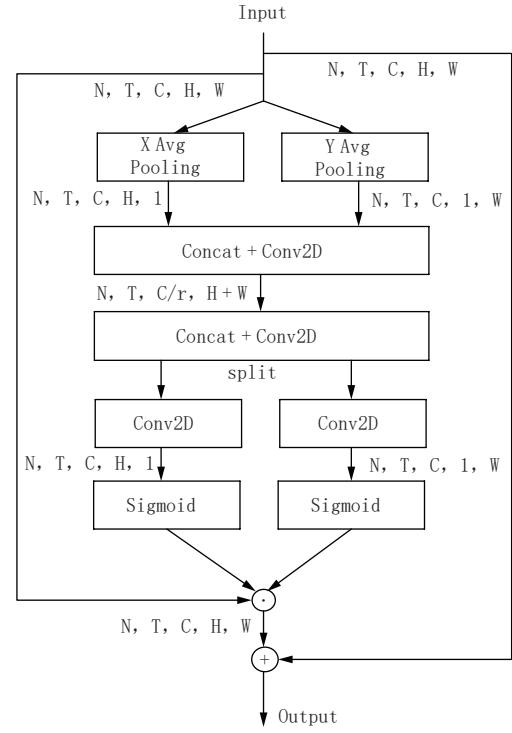


图 2 CA 模块<sup>[11]</sup>  
Fig. 2 The coordinate attention module

的数学表达式为:

$$L_{\text{SSIM}} = \text{SSIM}_{G, P} \circ \quad (13)$$

内容损失  $L_{\text{content}}$  计算目标图像  $G$  与预测图像  $P$  像素点之间的欧式距离, 其中为 L2 范数。内容损失可以定义为:

$$L_{\text{content}} = \|P - G\|_2 \circ \quad (14)$$

为实现结构损失函数与内容损失函数之间权值平衡, 通过赋予结构损失相应的超参数  $\lambda (= 2)$  提高模型的泛化能力, 综上 NGPNet 整体损失函数可表示为:

$$\text{Loss} = \lambda L_{\text{SSIM}} + L_{\text{content}} \circ \quad (15)$$

## 2 实验与结果分析

### 2.1 数据集

美国国立肺筛查试验 (National Lung Screening Trial, NLS) 是一项在美国进行的大型随机多中心肺癌相关研究, 在 NLST 数据集中登记的许多受试者涵盖了多个时间点的 CT 扫描。在本文研究中, 为获取肺结节生长趋势预测的研究数据, 我们从 NLST 中筛选出 312 名受试者, 共 786 次 CT 扫描, 其中共包含 423 例长期肺结节, 用于本文研究。扫描时间间隔为 3 到 24 个月, 每两个扫描时间点对同一个结节, 选取前后扫描样本组成样本对 (例如一个结节在扫描时间点  $t_1$ 、 $t_2$ 、 $t_3$  有 3 次随访扫描, 我们选择时间点  $t_1$  &  $t_2$ ,  $t_1$  &  $t_3$ ,  $t_2$  &  $t_3$  构成 3 个样本对), 共生成 926 例成对肺结节, 其中选取 700 对作为训练集, 150 对作为验证集, 76 对作为测试集。

### 2.2 实验预处理

对肺部三维 CT 影像进行预处理包括: 将 CT 扫描重新采样至  $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ , CT 影像的 HU 值截取到区间  $[-1200, 600]$ , 使用映射函数将取值标准化至  $[-1, 1]$  受试者的多次 CT 扫描进行人工图像配准。对配准的同一结节剪切出大小为  $48 \times 48 \times 48$  像素的立方体块。

### 2.3 实验环境及训练设置

实验软硬件环境配置由表 1 所示。训练过程中采用 Adam 作为优化器, 初始学习率为 0.001, 每 50 个 epoch 衰减到原始值的一半。将总训练 epoch 设置为 300, 批次大小为 16。

表 1 实验环境  
Table 1 The experimental environment

名称	配置
操作系统	Ubuntu16.04.1 LTS
编程语言	Python3.7.8
AI 框架	PyTorch 1.9.0
CPU	Intel(R) Core(TM) i7-7820X CPU @ 3.60GHz
GPU	NVIDIA GeForce GTX 1080Ti*2(11GB*2)
内存	32GB

### 2.4 评估指标

对生成的肺结节预测图像, 本文采用常用的峰值信噪比 (Peak signal-to-noise ratio, PSNR) 及 SSIM 指标进行图像质量评价。对 PSNR 通常表示的时候取 log 变成分贝 (dB), 其中 MSE 为均方误差, PSNR 定义式如下:

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right) = 20 \cdot \log_{10} \left( \frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \right), \quad (16)$$

其中 MSE 为最小均方误差,  $\text{MAX}_I$  表示图最大像素值, 通常, 如果像素值由  $B$  位二进制来表示, 那和  $\text{MAX}_I = 2^B - 1$ ,  $\text{MAX}_I = 2^B - 1$ 。

## 2.5 实验结果

本文将提出的算法与使用基准 3D U-Net 网络、采用形状变换分支、采用纹理学习分支及双分支网络生成的结果进行 PSNR、SSIM 指标分析, 相应的结果如表 2 所示。

同使用基准 3D U-Net 网络相比, 采用单形状学习分支 PSNR 提升 1.97 dB; 采用单纹理学习分支相比形状分支而言效果并不明显; 形状结合纹理双分支, 可以更好的进行运动变换和纹理细节的互补学习, 同基准相比提升了 3.19 dB, 证明了双分支结构的有效性。

通道注意力 SE<sup>[10]</sup> 是较为常用的注意力模块, 通过 SE 替换 CA 的对比实验, 来验证 CA 模块的有效性, 实验结果如表 3 所示。相较于会忽略位置信息的 SE 通道注意力机制, CA 更为注重空间位置信息, 这对于图像生成任务更为有效。

因相关研究的数据集并未公开, 本文模型同 NoFoNet<sup>[7]</sup> 采用类似数据处理方式的情况下, 进行结果对比分析。本文模型在未使用结节分割掩码的情况下, 同使用分割模掩码的 NoFoNet<sup>[7]</sup> 相比, 结果如表 4 所示, 本方法在减少标注结节分割掩码的前提下, PSNR 仍能获得较为一致的结果, 证明了本文方法与同类研究相比具有一定的优越性。

表 3 注意力对比实验

Table 3 The attention contrast experiment

方法	PSNR/dB	SSIM
SE	19.11	0.8479
CA	19.31	0.8513

表 2 实验结果

Table 2 The results of the experiment

方法	PSNR/dB	SSIM
基线 (3D U-Net)	16.12	0.8056
形状分支	18.09	0.8375
纹理分支	18.20	0.8398
形状 + 纹理双分支	19.31	0.8513

表 4 同类方法对比实验

Table 4 Comparative experiment of similar methods

方法	PSNR/dB
NoFoNet	18.21
本文	19.31

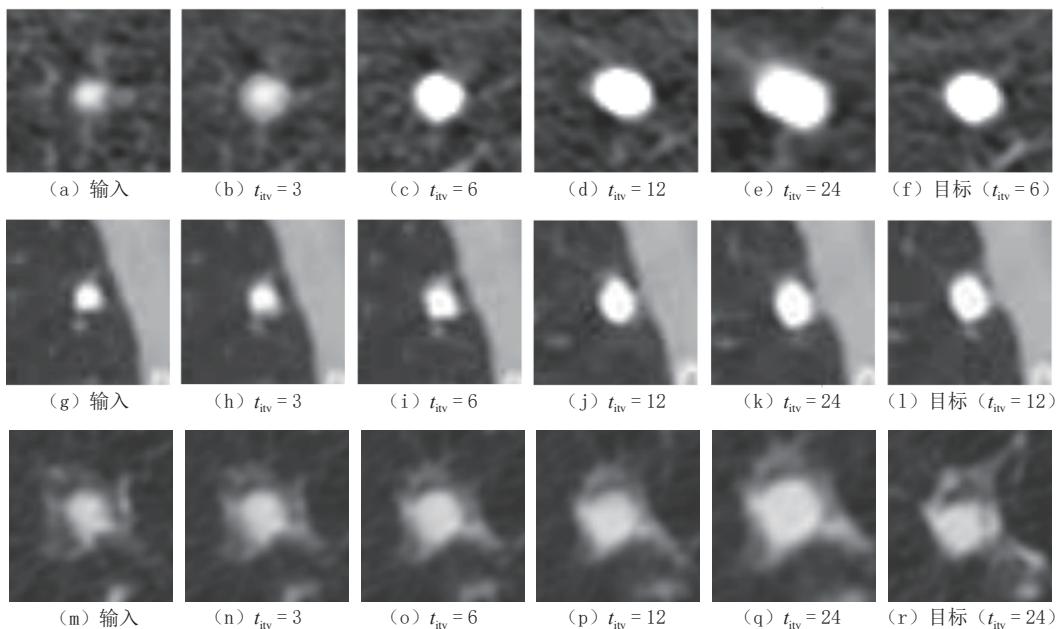


图 3 基于时间间隔预测的肺结节图像

Fig. 3 Lung nodule images based on time interval predictions

图3 对不同时间间隔 ( $t_{hv} = 3, 6, 12, 24$ ) 的预测图像进行展示, 由图可见预测的图像同真实目标图像相比, 具有较高的生成质量及准确性。

### 3 结论与展望

本文提出了一种结节生长预测网络模型, 该模型可以生成指定时间间隔肺结节的生长预测图像。通过图像质量评价指标及可视化的结果显示, 网络可生成高质量预测结果, 用来帮助医生做出更为准确的随访决策。

本文研究的数据量相对较小, 后续可收集更多时间间隔更长数据对扩充训练样本量, 或通过数据合成方法来进一步增大数据集, 从而提升模型的预测效果。后续可引入结节分割及密度分类信息进一步精细化控制肺结节生成效果。

### 参考文献

- [1] OUDKERK M, LIU S, HEUVELMANS M A, et al. Lung cancer LDCT screening and mortality reduction: Evidence, pitfalls and future perspectives[J]. *Nature Reviews (Clinical Oncology)*, 2021, 18(3): 135–151. DOI:[10.1038/s41571-020-00432-6](https://doi.org/10.1038/s41571-020-00432-6).
- [2] SHEN W, ZHOU M, YANG F, et al. Multi-scale Convolutional Neural Networks for lung nodule Classification[C]//Information Processing in Medical Imaging: 24th International Conference, UK: Springer, 2015: 588–599.
- [3] SHEN W, ZHOU M, YANG F, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification[J]. *Pattern Recognition*, 2017, 61: 663–673. DOI:[10.1016/j.patcog.2016.05.029](https://doi.org/10.1016/j.patcog.2016.05.029).
- [4] ZHANG L, LU L, SUMMERS R M, et al. Convolutional invasion and expansion networks for tumor growth prediction[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(2): 638–648.
- [5] RAFAEL-PALOU X, AUBANEL A, BONAVITA I, et al. Re-identification and growth detection of pulmonary nodules without image registration using 3D siamese neural networks[J]. *Medical Image Analysis*, 2021, 67: 101823. DOI:[10.1016/j.media.2020.101823](https://doi.org/10.1016/j.media.2020.101823).
- [6] SHENG J, LI Y, CAO G, et al. Modeling nodule growth via spatial transformation for follow-up prediction and diagnosis[C]//2021 International Joint Conference on Neural Networks (IJCNN). Shenzhen: IEEE, 2021: 1–7.
- [7] LI Y, YANG J, XU Y, et al. Learning tumor growth via follow-up volume prediction for lung nodules[C]//Proceedings of the 23th International Conference on Medical Image Computing and Computer-assisted Intervention. Peru: Springer, 2020: 508–517.
- [8] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention. Munich: Springer, 2015: 234–241.
- [9] BALAKRISHNAN G, ZHAO A, SABUNCU M R, et al. VoxelMorph: A learning framework for deformable medical image registration[J]. *IEEE Transactions on Medical Imaging*, 2019: 1788–1800. DOI:[10.1109/TMI.2019.2897538](https://doi.org/10.1109/TMI.2019.2897538).
- [10] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023. DOI:[10.1109/TPAMI.2019.2913372](https://doi.org/10.1109/TPAMI.2019.2913372).
- [11] 唐秉航, 王艳芳, 马力, 等. 基于混合注意力机制的肺结节假阳性降低[J]. *CT 理论与应用研究*, 2022, 31(1): 63–72. DOI:[10.15953/j.ctta.2021.002](https://doi.org/10.15953/j.ctta.2021.002).
- [12] TANG B H, WANG Y F, MA L, et al. False positive reduction of pulmonary nodules based on mixed attentional mechanism[J]. *CT Theory and Applications*, 2022, 31(1): 63–72. DOI:[10.15953/j.ctta.2021.002](https://doi.org/10.15953/j.ctta.2021.002). (in Chinese).
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770–778.
- [14] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612. DOI:[10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).

# Predicting Lung Nodule Growth with Shape Transformation and Texture Learning

MA Li<sup>1</sup>, HUANG Dehuang<sup>2</sup>, WANG Yanfang<sup>1✉</sup>

1. Zhongshan Yangshi Technology Co., Ltd, Zhongshan 528400, China

2. Zhongshan research Institute, Beijing Institute of technology, Zhongshan 528400, China

**Abstract:** While artificial intelligence has achieved considerable maturity in lung nodule detection, research on growth prediction remains limited. Accurate growth prediction aids clinical decision-making, informing patient follow-up strategies. This paper proposes a novel nodule growth prediction network model that generates high-quality lung nodule images at specific time intervals. The model employs a two-branch structure for feature extraction. One branch, leveraging a displacement field prediction mechanism, models the shape transformation of pulmonary nodules through voxel-level future displacement estimation. The other branch, empowered by a three-dimensional U-Net, focused on learning texture changes within the nodules. A coordinate attention mechanism that emphasizes informative features within the extracted high-dimensional feature map. Subsequently, the outputs of both branches are fused and fed into the feature reconstruction module to generate the final lung nodule growth prediction image. Furthermore, a time interval coding module is introduced to incorporate the desired time interval into the network, enabling the generation of prediction images for different future time points.

**Keywords:** lung nodules; growth prediction; displacement field; time interval coding



**作者简介:** 马力, 男, 中山仰视科技有限公司 CTO, 主要从事人工智能深度学习技术在医学影像上的开发与应用系列研究, E-mail: [ma20230704@163.com](mailto:ma20230704@163.com); 王艳芳✉, 女, 中山仰视科技有限公司 CEO, 主要从事人工智能深度学习技术在医学影像上的开发与应用系列研究, E-mail: [yfwang6@sina.cn](mailto:yfwang6@sina.cn)。