RADIFUSION：多影像组学深度学习乳腺癌风险预测模型，基于序列乳腺X线图像，并

Hong Hui Yeoh[1], Andrea Liew[1], Raphaël Phan[7], Fredrik Strand[3,4], Kartini Rahmat[5], Tuong LinhNguyen[6], John L. Hopper[6], Maxine Tan[1,2,*]

[1] 马来西亚莫纳什大学工程学院电气与计算机系统工程学科，Bandar Sunway 47500，马来西亚。

[2]俄克拉何马大学电气与计算机工程学院，诺曼，俄克拉何马州 73019，美国。

[3]卡罗琳斯卡学院肿瘤与病理学系，索尔纳，瑞典。

[4]卡罗林斯卡大学医院乳腺放射学科，索尔纳，瑞典。

[5]马来亚大学医学院生物医学成像系及马来亚大学研究成像中心，马来西亚吉隆坡 50603。

[6]墨尔本大学人口与全球卫生学院流行病学与生物统计学中心，澳大利亚维多利亚州墨尔本。

[7] 马来西亚莫纳什大学信息技术学院，马来西亚 Bandar Sunway 47500

[*]通讯作者。电子邮箱：Maxine.Tan@monash.edu（Maxine Tan）

摘要

乳腺癌是一个重要的公共卫生问题，早期检测对于高风险患者的分诊至关重要。序列筛查乳腺X线摄影图像能够提供有关乳腺组织随时间变化的重要时空信息。本研究提出一种名为 RADIFUSION 的深度学习架构，该架构利用序列乳腺X线图像，并结合线性图像注意力机制、影像组学特征、用于融合不同乳腺摄影视图的新型门控机制，以及基于双侧不对称性的微调，以评估乳腺癌风险。我们使用名为筛查年龄女性队列（CSAW）的筛查数据集评估模型，该数据集共包含 8,723 名患者。根据由 1,749 名女性组成的独立测试集所得结果，我们的方法优于其他最先进模型；在 1年AUC、2年AUC 和 > 2年AUC 这三项指标上，受试者工作特征曲线下面积（AUC）分别为 0.905、0.872 和 0.866。本研究强调，融入图像注意力、影像组学特征、门控机制和基于双侧不对称性的微调等多种深度学习机制，有助于提高乳腺癌风险评估的准确性。我们还证明，利用序列乳腺X线图像中的时空信息能够提升模型性能。研究结果表明，RADIFUSION 可为临床医生提供一种强有力的乳腺癌风险评估工具。

关键词

癌症风险预测，乳腺摄影，自注意力，卷积神经网络，深度学习，影像组学

1. 引言

乳腺癌是女性患者中最常确诊的癌症，也是美国癌症相关死亡的第二大原因 [1]。乳腺癌占所有女性癌症的30%，并使20至49岁女性在成年早期的癌症风险增加77%。乳腺癌的早期检测至关重要，因为这能显著提高治疗成功的机会和患者生存率。因此，定期进行筛查检查，例如乳腺X线影像和临床乳房检查，有助于在乳腺癌早期发现病症。

除了早期检测之外，风险评估在乳腺癌的预防和管理中也十分重要。这包括识别和评估可能增加个人风险的各种因素

，例如年龄、家族史、遗传因素和生活方式因素。早期风险模型（如Gail模型）[2] 和Tyrer-Cuzick模型[3] 利用了这些因素，但性能有限。乳腺密度和实质纹理模式也被认为与乳腺癌风险增加相关，对其进行评估也已证明有助于风险预测[4-6]。然而，我们认为，深度学习模型还可以同时从乳腺X线摄影图像中捕获乳腺密度或实质纹理模式等上下文信息。深度学习已被用于开发准确的乳腺癌风险评估模型。基于图像的深度学习模型[7-9] 已显示出非常有前景的结果，表明深度学习有能力从乳腺X线摄影图像中识别出比传统风险模型更多的预测性信息。

据我们所知，除一项研究外，所有基于深度学习的研究[10]都未在模型中考虑此前乳腺X线筛查所提供的信息。在标准临床实践中，女性会接受多次筛查，以监测致密乳腺组织及其外观随时间发生的变化。其他任务中已有研究提出采用乳腺X线影像开展纵向研究[11-13] ，并表明纳入既往筛查信息可以提升模型的预测能力。一项采用门控循环单元（GRUs）[10] 进行乳腺癌风险预测的新研究也显示出类似的改进，尽管该研究未考虑影像组学特征或任何注意力机制。

尽管已有许多研究提出通过临床风险因素[2, 3],、乳腺实质纹理分析／影像组学方法[14], 、基于当前年度乳腺X线摄影图像的深度学习[6], 或基于双侧不对称性的方法[15] 来预测癌症风险，但这些研究均采用单一模型，未结合所有因素构建全面的风险模型。另一方面，影像组学的各个方面、左右乳房之间的双侧不对称性（因为癌症通常只发生在一侧乳房，而正常乳房通常外观对称）、深度学习所学习到的特征以及临床风险因素（如年龄、体质指数、家族史等），均已分别被证明能够为风险预测提供有用信息。现有文献缺少的是一种将所有这些因素整合到一个全面模型中的风险模型。

非局部网络最初用于视频分类 [16] ，其目的是通过对整个视频中生成的所有特征进行加权求和，将卷积神经网络（CNN）的感受野扩展至输入图像大小。对于高分辨率乳腺X线图像，像非局部网络这样的自注意力机制方法非常适合提取乳腺中的细微模式/组织特征，以指示乳腺癌风险的存在。与 VIT 等其他基于 Transformer 的方法相比，非局部网络的另一优势是可用于三维视频 [17] 。这对我们工作的意义在于，我们可以将乳腺X线影像的时间序列建模为三维视频。这样，便可以从同一患者的一系列序列乳腺X线图像中检测癌症风险（以检测患者乳腺组织中的细微变化），这与放射科医师在临床实践中的做法类似。然而，Transformer/自注意力模型的缺点是，其复杂度与输入序列长度的平方成正比，因为需要计算每对位置上输入表示之间的点积 [18]。因此，标准的 Transformer/自注意力模型难以高效处理较长的输入序列，更不用说大型二维图像（乳腺X线影像）和三维视频了。

在本研究中，我们提出了一种名为 RADIFUSION 的新型深度学习架构，通过对乳腺组织结构随时间发生的变化进行建模，实现乳腺癌风险预测。我们提出的深度学习模型融入了新方法，并将此前研究过的风险因素纳入其中或加以扩展，包括深度学习特征、影像组学、双侧不对称特征检测，以及整合来自不同头尾位（CC）和内外斜位（MLO）视图的特征等，从而构建一个全面的风险模型。首先，我们提出了一种新型注意力机制，能够高效应对三维序贯筛查乳腺X线摄影图像带来的计算和内存需求增长。据我们所知，这是首个用于图像/视频的加性注意力模型。这种新型注意力机制使模型能够快速、准确地聚焦于患者多次既往筛查中最相关的特征。其次，我们提出了新的影像组学特征，以补充深度学习特征，增强模型的风险预测能力。第三，

我们提出了一种门控机制，可自动确定应赋予（头尾位或内外斜位）乳腺影像视图的最佳权重/注意力，以提升网络性能。最后，我们提出了一种新的伪标签方法，利用从双侧不对称性中提取的信息进一步微调并提升模型性能。

作者希望我们的新模型能够帮助放射科医师将高风险患者与低风险患者区分开来。我们还希望新模型能够协助放射科医师制定个体化筛查建议，以便对短期内（例如 1 至 2 年）癌症易感性较高的女性比短期风险较低的女性进行更频繁的筛查。我们在一个全面的筛查数据集——筛查年龄女性队列（CSAW）数据集 [19] 上评估了模型的性能。

本研究的成果包括：

1. 一种名为 RADIFUSION 的新型深度学习架构，可通过序列乳腺X线筛查捕捉乳腺组织随时空变化的特征。2. 一种大数据方法，将深度学习特征与影像组学特征相结合，并融合头尾位（CC）和内外斜位（MLO）视图的特征。3. 一种名为 SHIFT（空间通道图像 Fastformer）的新型注意力模块，能够高效捕捉多次乳腺X线筛查中的空间注意力和通道注意力。4. 一种新的多视图门控机制，可对头尾位（CC）和内外斜位（MLO）视图的信息进行最优加权。5. 一种新的双侧对称性感知微调方法，利用对侧乳房之间的双侧不对称风险评分进一步优化模型。6. 在全面的筛查数据集上验证新的风险模型及其性能。

2. 材料与方法 2.1 数据集

我们在一个名为筛查年龄女性队列（CSAW）数据集的综合筛查数据集上分析了我们提出的新方法 [19]。对于 CSAW 数据集，斯德哥尔摩伦理审查委员会批准了该研究，并豁免了获取个人知情同意的要求（EPN 2016/2600-31）；相关补充内容经瑞典伦理审查局批准（EPM 2019-01946、EPM 2019-03638、EPM 2021-01030）。

所获取的数据集是 CSAW [19]，的一个子集，该数据集是一个基于全体人群的完整女性队列，纳入了 2008 年至 2016 年期间受邀在瑞典斯德哥尔摩地区接受筛查的 40 至 74 岁女性。每隔 18 至 24 个月，对年龄为 40 至 74 岁的女性采集多个视图的乳腺X线影像。所有图像均使用 Hologic 乳腺摄影设备采集。对于检查未包含全部四个标准视图，或年龄超出 40-74 岁筛查年龄范围的患者，均予以排除。

所提供的数据集共包含 8,723 名患者，其中对照组 7,850 名，病例 873 名。对于这 873 名病例，数据集对两个关键临床风险因素进行了删失处理，即患者年龄以及从筛查到癌症诊断的天数。对于患者年龄，在乳腺摄影时年龄为 40–55 岁的患者被赋予类别/值 1，年龄为 **55+** 岁的患者被赋予类别/值 2。对于确诊癌症的患者，在筛查后 **<= 60** 天确诊的患者被赋予类别 1，61–729 天确诊的患者被赋予类别 2，**730+** 天确诊的患者被赋予类别 3。出于匿名化目的，每位癌症患者因此被归入宽泛类别，这可能限制模型从临床风险因素中学习的能力。在这 873 例归入上述 3 个宽泛类别的病例中，一些患者截至当前年份有序列乳腺X线图像（最多 5 次），而另一些患者则只有一张筛查乳房X线摄影图像，放射科医生在该图像中检测到癌症。表 1 列出了这 873 例第 1 至第 3 类病例中每位患者可用的连续筛查次数分布，图1对此进行了展示。图 2 展示了数据集中一名具有三次纵向既往乳腺X线检查的癌症患者示例。

表 1. CSAW 数据集按患者年龄以及从筛查到癌症诊断的天数划分

	亚组	阳性	阴性
Age	40 - 55	319 (37%)	3853 (49%)
	55+	554（63%）	3997（51%）
癌症诊断	0	-	7850
	1（<= 60 天）	524（60%）	-
	2（61-729天）	217（25%）	-
	3（730+ 天）	132（15%）	-

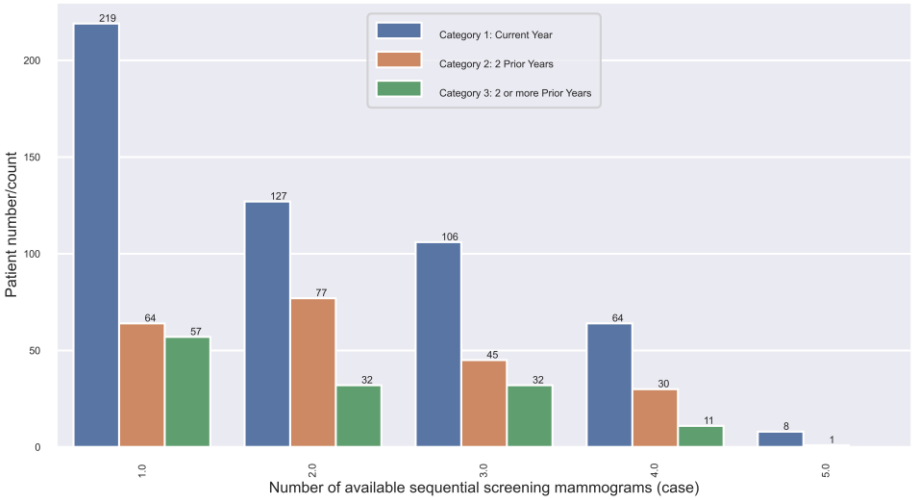


图1. CSAW 数据集中每位患者在当前及既往三种癌症诊断类别中的可用连续筛查次数。

2.2 数据预处理 我们应用了一些预处理

将乳腺X线影像输入我们的风险模型之前，我们对其进行了以下处理。首先，我们通过大津阈值法分割并去除了无关标签和背景。所有乳腺摄影视图均调整为 256×256 像素，并采用零填充以保持图像的纵横比。我们还通过计算每个交叉验证折训练集中的像素均值和标准差，将图像归一化为均值为零、方差为一。

为了利用既往乳腺X线筛查图像，我们构建了乳腺X线摄影图像的三维表示，将处理后的乳腺X线影像按筛查时间从近到远依次堆叠。这可以理解为乳腺X线影像的视频形式，不同筛查对应视频中的不同时间点。对于筛查次数不足的患者，我们复制其最近一次筛查的图像，以填补视频中的时间点空缺。

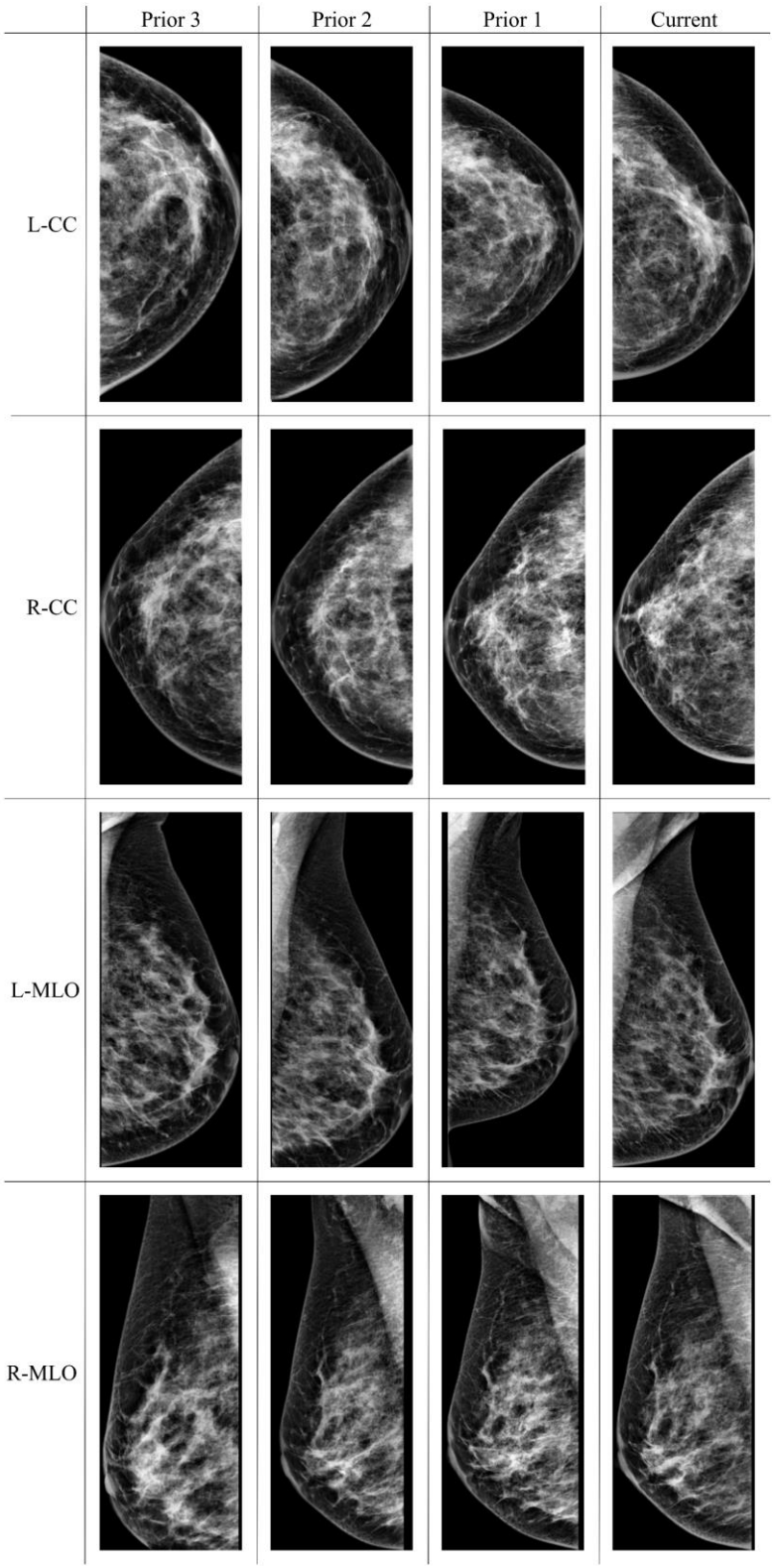


图2. 我们数据集中一名癌症患者的三次纵向既往乳腺X线检查示例。标记为prior-1至prior-3的既往检查采集于2009年至当前检查2013年期间。每次既往检查包含四幅图像，分别为左右乳房的头尾位和内外斜位。

2.3 三维卷积神经网络（CNN）架构 我们采用Resnet18作为CNN模型的骨干网络。它作为特征提取器，将图像表示压缩为一个维度为512的向量，如图5所示。该

基础架构（该架构用于图像分类）被扩展为三维（3D）模型，方法是“扩展”卷积核[16, 20]。

我们的方法采用自[16], ，其中二维 k 卷积核可以扩展为三维 t k 卷积核，使其覆盖 t 帧。使用预训练二维 ResNet-18 模型的权重来初始化卷积核，其中 t k 卷积核中的每个 t 平面都使用预训练的 k 权重进行初始化，并按 $1/t$ 重新缩放。这种初始化设置产生的结果与二维预训练模型在单个静态帧上运行并在时间域中重复该帧时的结果相同。

在我们的设置中，每个 ResNet 层中的第一个卷积核都使用3 3 卷积核进行初始化。其余卷积核则通过将其初始化为1 3 卷积核，简单地扩展为执行相同操作的三维对应卷积核。这种方法不会显著增加参数数量，从而缓解过拟合。

2.4 空间通道图像 Fastformer（SHIFT）模块 本节将介绍我们新提出的图像/视频注意力

模块 SHIFT（空间通道图像 Fastformer），并描述其运行方式。在自注意力中，使用点积注意力机制，通过相似度度量对查询与键值之间的交互进行全面建模，以确定哪些键区域与查询区域最为相似。这一过程的复杂度需要耗费大量资源，因此对于较大的输入，尤其是三维输入而言，实际应用相当困难。在我们的应用中，我们使用的是筛查乳腺X线摄影图像的三维序列，并将时间维度（即筛查序列）视为视频中的第三维度。随着序列筛查乳腺X线摄影图像数量的增加，输入规模不断扩大，注意力机制的内存开销和计算复杂度呈指数级增长。

针对经典自注意力的关键缺陷，本文提出了一种基于图像/视频的快速注意力机制，该机制受 Fastformer [21]中针对文本建模所提出的注意力机制启发，并采用加性注意力。据我们所知，这是首次针对图像/视频，尤其是医学图像，研究加性注意力模型。此外，在基于文本的 Fastformer [21], 中，注意力机制仅作用于不同词语之间。我们提出的图像 Fastformer 将加性注意力扩展到了空间注意力和通道注意力机制。根据我们先前的研究 [22, 23] 以及相关文献（例如挤压与激励网络、双重注意力网络 [24, 25]），可以观察到，通道重要性也是提升网络性能的重要组成部分。

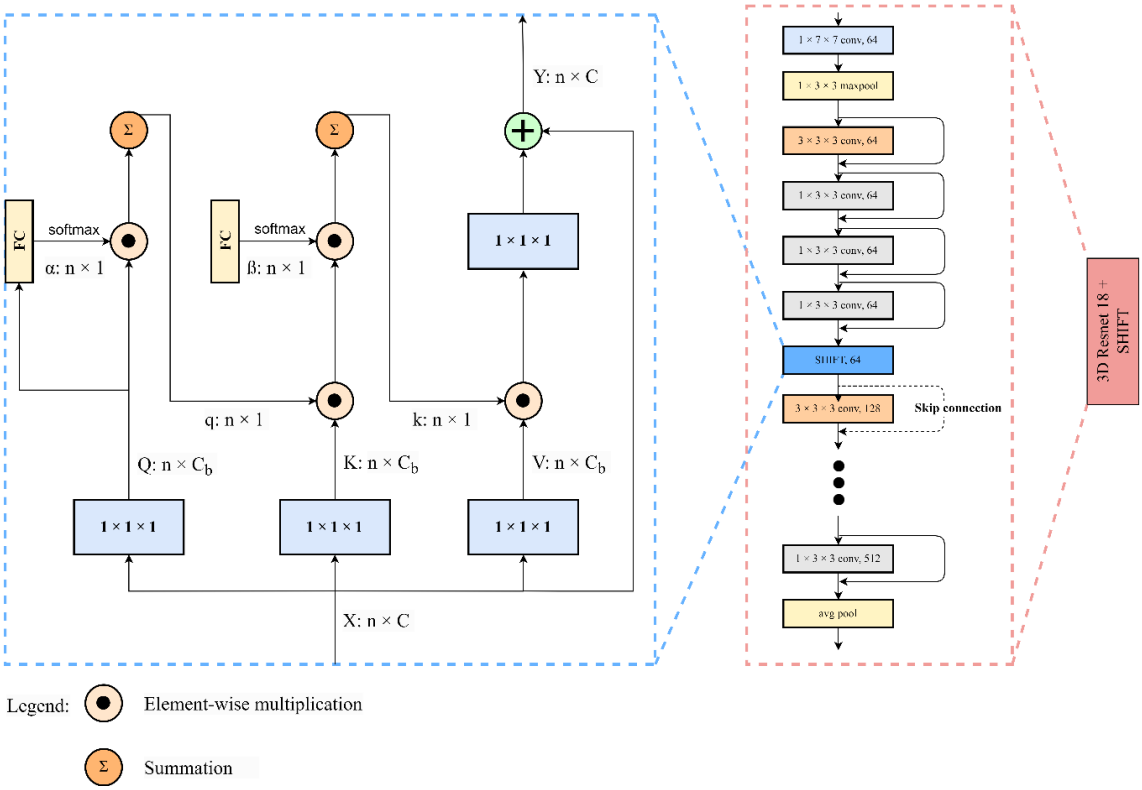


图3. 我们新提出的 SHIFT 模块框图及其在 3D ResNet-18网络中的位置。

与非局部网络类似，[16], 我们利用三个不同的线性层，将给定的输入$X \in \mathrm{R}$　　转换为查询、键和值矩阵Q、K、$V \in \mathrm{R}$　　其中，　　为了在查询矩阵中忽略编码上下文，我们首先计算标量值，α 并利用这些标量值计算重要区域的加权和：

$$\alpha = softmax\left(FC_Q(Q)\right) \in \mathbb{R}^{n \times 1} \tag{1}$$

其中，$FC_Q \in$　　是一个全连接层。α 的值即为注意力权重，负责识别模型应重点关注的多张（连续）乳腺X线影像中的关键区域。此步骤还避免了创建大小为　　、其规模/复杂度呈指数级增长的注意力矩阵的需要。随后计算全局注意力查询，如下所示：

$$q = \sum_{j=1}^{C_b} \alpha Q_j \in \mathbb{R}^{n \times 1}, \tag{2}$$

它概括了查询中的编码上下文。沿通道维度求和是直观的，而沿体素维度求和则不同，因为后者会丢失特征图中的关键信息。这与原始 Fastforme　　）构造其全局问题。

随后，我们通过逐元素乘积将这些全局信息传递给键，并对它们之间的交互进行建模，从而生成全局上下文感知键矩阵 p:

$$p = q * K \in \mathbb{R}^{n \times C_b} \tag{3}$$

类似于　对于查询，我们使用全局上下文感知键矩阵p计算注意力权重 ß:

$$ß = softmax\left(FC_K(p)\right) \in \mathbb{R}^{n \times 1}, \tag{4}$$

其中 $FC_K \in \mathrm{R}$ 是一个全连接层。它封装了相关的查询-键交互，随后这些交互将传递到值 FC_K 都通过将每个通道用作所有像素位置的注意力值/权重，高效地处理通道维度。通过加权和组合这些特征，可以生成一个全局查询/键向量。

使用两个全连接层对各通道的重要性进行加权也很重要，因为将 FC_K 替换为 FC_Q 会降低注意力机制的性能。我们推测，这两个全连接层负责汇总查询和查询-键交互中的不同信息。

全局键可使用注意力权重 ß 按如下方式计算：

$$k = \sum_{j=1}^{C_b} ßK_j \in \mathbb{R}^{n \times 1} \tag{5}$$

接下来，与查询-键交互类似，我们也通过对 k 和 V进行逐元素乘法来建模键-值交互：

$$u = k * V \in \mathbb{R}^{n \times C_b} \tag{6}$$

这样，值向量便可与全局键和全局查询向量进行交互。最后，我们将值矩阵 u 输入线性变换层，以获得与输入相同的通道维度。随后，将注意力特征加回图像特征，形成最终输出。根据实验观察，原始 Fastformer 中的查询添加是冗余的，因此予以省略。我们将这一模块统称为 SHIFT（空间通道图像 Fastformer）模块。

2.5 用于深度学习的影像组学融合 影像组学是提取 q 的过程

利用计算机算法从医学图像（如乳腺X线摄影图像）中提取定量特征 [26]。这些特征能够提供有关组织结构的补充信息，可用于诊断、预后和治疗规划。在乳腺摄影中，影像组学特征可用于区分恶性病变和良性病变 [27, 28] ，并预测乳腺癌风险 [15, 29]。

许多研究表明，影像组学能够通过提供有关组织结构和异质性的补充信息，增强传统乳腺摄影的效用 [4]。将深度学习特征与影像组学特征相结合，似乎是改进乳腺癌风险预测的自然发展方向。

影像组学的一项关键优势在于无需收集额外数据，因为它可以从现有的乳腺摄影数据中生成。与收集年龄、生育史、基因突变、激素因素、家族史等其他风险因素相比，这种方式更加便利。此外，这些数据往往缺失，也不易汇总。使用开源 Python 库 Pyradiomics [30]也可以轻松提取影像组学特征。我们研究中提取的大多数特征均通过 Pyradiomics [30] 库获得；使用开源库有助于复现我们所述的方法。表 2 展示了我们用于风险预测的 Pyradiomics 工具箱中的特征组/类型及其对应的数量和说明。

表 2. 根据分组或类型，从 Pyradiomics 库计算的影像组学特征及其描述。

特征组/类型	特征 数量/计数	描述
一阶	18	能量、总能量、熵、最小值、10th 百分位数、90th 百分位数、最大值、均值、中位数， 四分位距、极差、平均绝对偏差，

		稳健 平均绝对偏差、均方根、 偏度、峰度、方差、均匀性
灰度共生矩- 现矩阵 （GLCM）	23	自相关、联合平均、簇突出度、 簇阴影、簇趋势、对比度、相关性、 差异平均值、差异熵、差异 方差、联合能量、联合熵、信息相关性 度量1、信息相关性度量 2、逆差矩、最大 相关系数、逆差矩 归一化，逆差，逆差 归一化，逆方差，最大概率， 和熵，平方和
灰度大小 区域矩阵 （GLSZM）	16	小区域强调，大区域强调，灰度级 非均匀性，灰度级非均匀性 归一化，大小区域不均匀性，大小区域 非均匀性归一化，灰度级方差， 区域方差，区域熵，低灰度级区域 强调，高灰度区域强调，小区域 低灰度级强调，小区域高灰度级 强调，大区域低灰度强调，大 面积高灰度级强调
灰度游程 长度矩阵 （灰度游程长 度矩阵）	14	短游程强调、长游程强调、灰度级 非均匀性、灰度级非均匀性 归一化、游程长度非均匀性、游程长度 非均匀性归一化、游程百分比、灰度 级方差、游程方差、游程熵、低灰度 级游程强调、高灰度级游程强调、 短游程低灰度级强调，短游程高 灰度级强调，长游程低灰度级 强调，长游程高灰度级强调
邻域灰度 差 矩阵（NGTDM）	5	粗糙度、对比度、忙碌度、复杂度、强度
灰度级 依赖矩阵 （GLDM）	15	小依赖强调、大依赖 强调、灰度级非均匀性、依赖 非均匀性， 依赖非均匀性归一化，灰度级 方差， 依赖方差、依赖熵、低灰度 灰度级强调、高灰度级强调、小 依赖低灰度级强调、小 依赖高灰度级强调，大 依赖低灰度级强调， 大依赖高灰度级强调

除了标准影像组学特征外，我们还添加了新特征，以提升模型的能力：补充特征包括使用离散余弦变换（DCT）和快速傅里叶变换（FFT）从频域中的两张图像映射计算得出的统计特征。采用这些特征已被证明能够改善乳腺癌良恶性分类[27]，并且也应能增强风险评估。

设$f(x, y)$为大小为M 乘N的输入图像，二维离散余弦变换可表示为：

$$DCT(m,n) = \frac{2}{\sqrt{MN}} C(m)C(n) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cos(\frac{(2x+1)m\pi}{2M}) \cos(\frac{(2y+1)n\pi}{2N}), \tag{7}$$

其中$C($ [illegible] n [illegible] $C($ [illegible] $)$, [illegible] 然后，二维快速傅里叶变换可表示为：

$$F(m,n) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \exp(-j2\pi(\frac{xm}{M} + \frac{yn}{N})), \tag{8}$$

其中 [illegible] 和 [illegible] 。

由于可以从乳腺中脂肪组织和纤维腺体组织的分布模式中推断出关键信息，因此进一步分析乳腺X线摄影图像中的实质纹理十分重要。乳腺X线摄影图像局部的频谱或频率内容能够揭示特定模式。乳腺X线摄影图像的大部分信息通常集中在低频分量中，而噪声往往出现在高频中。根据计算得到的两种离散余弦变换和FFT频域图像映射，我们计算了表3所列的统计特征。

表3. 用于癌症风险预测的新型离散余弦变换和FFT频域统计特征。

特征组/类型	特征数量/计数	描述
离散余弦变换（DCT） 快速傅里叶变换 (FFT)	30	均值、最大值、方差、偏度、峰度、熵、能量、均方根、均匀性、最小值、中位数、极差、四分位距、平均绝对值偏差、中位数绝对偏差

先前研究表明，加入临床风险因素（例如女性年龄、家族史、主观乳腺密度和 BIRADS 分级）有助于提升风险模型的性能 [2, 。29]
因此，在我们的研究中，我们将患者年龄纳入模型；由于伦理/隐私限制，我们无法获取其他临床因素。

我们将所有计算得到的影像组学特征与患者年龄一并纳入模型的全连接层。我们最初尝试了文献中的多种方法来整合影像组学特征和临床特征，包括序列浮动前向选择 (SFFS)、序列浮动后向选择 (SFBS)、动态仿射特征图变换 (DAFT) [31], 等。然而，实验结果表明，将所有特征直接纳入全连接层效果最佳。这种轻量级方法十分有效，而且可以在全连接层中轻松调整相关/有用特征的权重，同时只需训练少量额外参数。

2.6 多视图门控机制 文献中，不同方法

和策略已被提出，用于结合来自四种不同乳腺摄影视图的图像特征 [32-34]。几乎所有已提出的方法都相当复杂，需要使用全连接层，涉及大量参数调优，和/或计算开销较大。本研究与参数较少且高效的 SHIFT 方法一致，考察了几种轻量但有效的不同视图融合方法。

在最初的方法中，我们采用最简单、最轻量的方式，对每个视图对应的全连接层的 4 个输出取平均。随后，模型的输出按如下方式计算：

$$y = \sigma\left(\frac{X_{LCC} + X_{RCC} + X_{LMLO} + X_{RMLO}}{4}\right), \tag{9}$$

其中，σ 表示 Sigmoid 函数，X_{LCC}、X_{RCC}、X_{LMLO}、X_{RMLO} 分别表示每个视图对应的全连接层输出，y 表示模型的输出。然而，这种方法赋予头尾位和内外斜位视图同等的重要性。相反，文献中的大多数研究结果（包括我们此前的研究）表明，在预测/检测乳腺癌时，某些视图比其他视图能提供更多信息 [9, 32, 35]。

为重新调整各视图的重要性，我们加入一个门控，以便模型判断哪个视图对风险评估的贡献更大。这使模型能够随机决定是限制还是增强来自各视图的信息。模型输出的计算方式相应更新如下：

$$y = \sigma\left(\frac{W_T X_{LCC} + W_T X_{RCC} + W_S X_{LMLO} + W_S X_{RMLO}}{2W_T + 2W_S}\right), \tag{10}$$

$$W_T, W_S = W_f + \tanh(W_\theta), \tag{11}$$

其中，W_T 和W_S 分别表示头尾位和MLO视图的缩放因子/门控。在两个缩放因子中，W_f 是固定权重，而W_θ 是可训练参数。设置固定权重W_f 是为了防止任一缩放因子在训练过程中降至0，从而彻底丢失该视图的信息。两个缩放因子最初均初始化为1，以模拟4个视图的平均设置；参见（9）。

2.7 基于双侧不对称性的微调 乳房之间的双侧不对称性通常

表明癌症发展，因为癌症通常只发生在一侧乳房[36]。我们此前的研究[15, 29, 32] 证实了这一点，并表明，基于双侧乳腺X线特征不对称性的模型所计算的风险评分，有望预测女性短期内罹患乳腺癌的风险。因此，我们首先暂时移除门控（见第2.6节），并为每个独立的乳腺影像视图生成风险评分。然后，我们利用这些评分计算对侧乳房之间的风险评分差异，如下所示：

$$y_{LCC} = \sigma(X_{LCC}), \quad y_{RCC} = \sigma(X_{RCC}), \quad y_{LMLO} = \sigma(X_{LMLO}), \quad y_{RMLO} = \sigma(X_{RMLO}), \tag{12}$$

$$\gamma = |(y_{LCC} + y_{LMLO}) - (y_{RCC} + y_{RMLO})|, \tag{13}$$

其中，y_{LCC}、y_{RCC}、y_{LMLO}、y_{RMLO} 分别表示一名患者左、右乳房的头尾位和内外斜位视图的风险评分，γ 表示左右乳房之间的不对称风险评分。新的不对称风险评分γ 表明哪些患者的乳腺摄影图像中存在双侧不对称性的迹象。

我们首先在训练子集上使用该方法训练模型。随后，我们采用伪标签方法，使用训练好的模型对训练集中的所有图像进行预测，并生成软标签及其对应的不对称风险评分。我们重新训练分类器，使其拟合训练集上获得的软伪标签。在实验中，我们观察到，病例训练集中的病例 的不对称风险评分平均高于对照组，这与我们此 致，证实病例中的双侧不对称性高于对照组（因为癌症通常仅发生在一侧乳房）。

在我们的方法背景下，可以将序贯筛查 CSAW 数据集表述如下：

$$D_s = \{(x_i, y_i, \bar{y}_i, \gamma_i)\}_{i=1}^{n}, \tag{14}$$

其中（x_i、y_i、$\bar{y}_i$、γ_i）分别表示数据集中i^{th} 患者的图像、真实标签、软标签和不对称评分；n 表示数据集中的患者数量。

CSAW 筛查数据集存在严重不平衡问题，共包含873例病例和7850例对照组，这可能会影响我们模型的训练/学习能力。由于数据集不平衡，我们尝试减少数据集中对照组的数量。具体而言，我们移除了不对称风险评分较高的患者，γ 方法是判断该评分是否超过P_T 训练集对照组中的某个百分位数，阈值百分比T 为可调超参数。这样可以消除模型判定为不一致的对照组（即高度不对称/假阳性），从而构建筛选后的训练集。这使我们能够纳入双侧不对称感知，以创建更理想的数据集，从而实现更好的微调。数据集中的病例得以保留，因为其占比本已很低，进一步减少病例数量会削弱模型的训练性能。因此，我们获得筛选后的训练数据集如下：

$$D_f \subset D_s = \{(x_i, y_i, \bar{y}_i, \gamma_i) \mid \big((\gamma_i < P_T) \cap (y_i = 0)\big) \cup (y_i = 1)\}_{i=1}^{m}, \tag{15}$$

其中，m 表示筛选后数据集中的患者总数，$m < n$。

通过实验，我们发现，仅在数据集D_s 或D_f 上进行训练会产生较差的结果。由于训练样本较少，仅在筛选后数据集上进行训练会导致模型性能较差；而仅在原始数据集上进行训练也不理想，因为其中包含具有较高双侧不对称性风险评分的对照组。为解决这一问题，我们改为先在数据集D_s 上对模型进行预训练，训练轮数为总训练轮数的一半，然后在数据集D_f上对模型进行微调。这样可以使模型充分利用两个数据集，从而提升其性能。整个过程如图4所示。

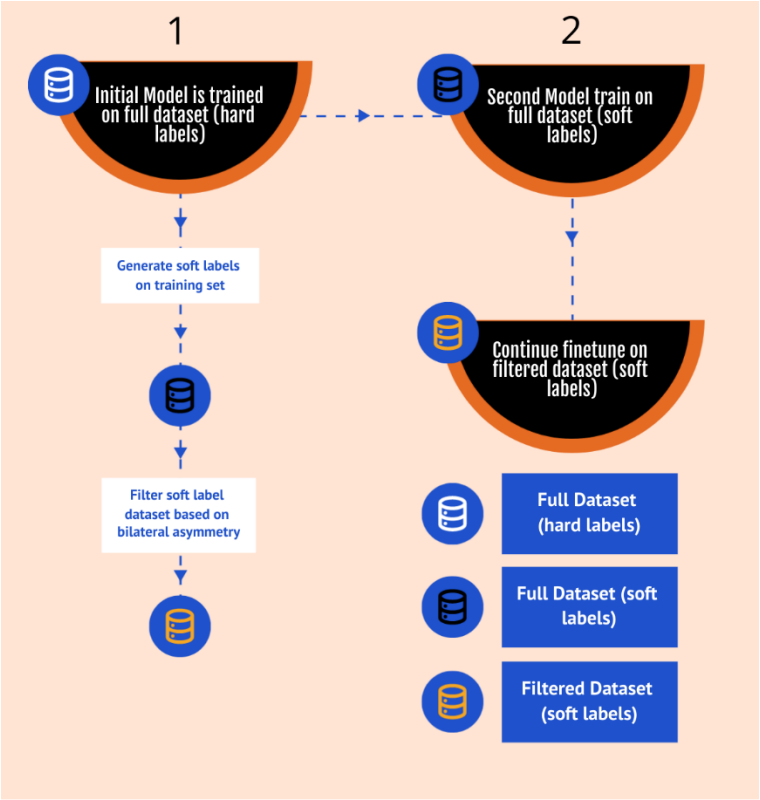


图4. 基于双侧不对称性的微调流程图。

2.8 所提出的架构 我们的框图

提出的深度学习架构如错误！未找到引用源。5所示。我们使用二元交叉熵损失训练网络，共训练60轮，批次大小为12（患者数量）。我们使用 Adam优化器 [37] 训练模型，学习率为0.0001。训练在两块 NVIDIA RTX 2080 Ti GPU 上进行。

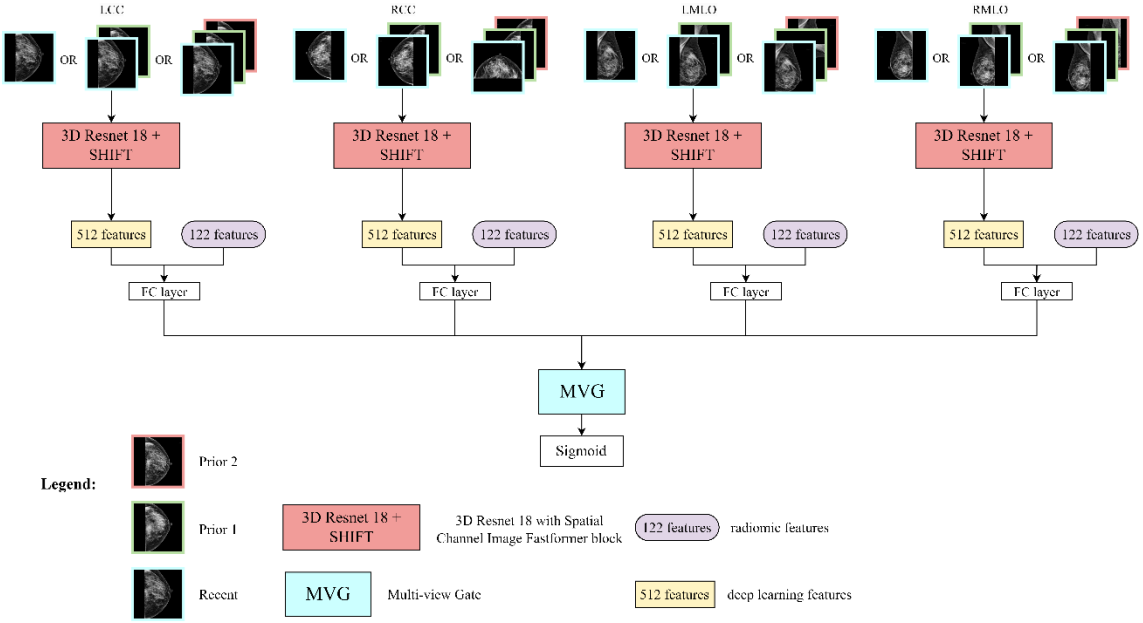


图5. 所提出的深度学习架构通过 SHIFT 模块捕捉乳腺组织结构随时间的变化，以预测乳腺癌风险，并支持不同数量的筛查（1 至 3 次）。模型接收来自 4 个视角的乳腺X线影像，以及输入全连接层的影像组学特征（122 个新增和既有特征）。模型采用了新的门控机制和一种基于双侧不对称性的微调方法，以输出最终风险评分。3D ResNet 18 + SHIFT 和全连接层的权重在所有 4 个视角间共享。

2.9 实验设置与分类方法

我们从数据集中随机抽取 20% 作为独立测试集，其中包含来自 1,749 名女性的 19,652 张乳腺X线影像。对数据集其余部分（包含来自 6,974 名女性的 79,136 张乳腺X线影像）进行按患者分层的五折交叉验证（CV）。每次交叉验证中，使用训练集（4 折）优化网络参数，剩余一折作为验证集。为确保数据集划分的一致性，我们尽可能在训练集、验证集和测试集中维持病例与对照组的比例。

按照[8], 中相同的评估方法计算2年AUC结果：如果病例在前一年的乳腺X线影像之后2年内被诊断出癌症，则视为阳性。我们采用类似的评估方法计算1年AUC和> 2年AUC结果。因此，我们将匿名化的真实标注结果评估为1年AUC、2年AUC和> 2年AUC，这些结果对应于数据集中提供的真实标签。所有结果均列于结果部分的相应表格中。我们采用自助抽样法计算所有AUC值的95%置信区间（CI）。所有脚本和卷积神经网络架构均使用 Python 软件包（v3.8.5）、Scikit-learn（v1.0.2）、Matplotlib（v3.5.1）和 PyTorch（v1.10.1）库实现。

3. 结果 表4表列

测试并比较了在基线3D ResNet-18模型上使用不同数量的乳腺X线筛查的结果。此处使用的模型是基线ResNet-18架构，患者风险评分取自4个视图的平均值。从结果可以看出，为模型提供额外的筛查乳腺 X 线摄影图像有助于更好地预测癌症风险。具体而言，筛查次数从1次增加到3次时，曲线下面积逐渐增大。这表明，3D卷积核能够捕捉序列乳腺X线图像之间的时空关系，从而取得整体更好的结果。

表4。CSAW 数据集中确诊癌症的3类患者在基线三维（3D）Resnet-18网络上的列表结果，对应1年AUC、2年AUC和> 2年AUC，并附有95%置信区间（CI）。该表还列出了

通过以下方式获得的结果：从数据集中选取不同数量的乳腺X线筛查（即1、2或3次筛查）。

模型	筛查次数筛查	1年曲线下面积（95%置信区间）	2年曲线下面积（95% CI)	> 2年曲线下面积（95% CI)
基线	3	0.882（0.853-0.913）	0.855 (0.829-0.886)	0.849 (0.823-0.878)
	2	0.879（0.85-0.91）	0.839（0.811-0.871）	0.835（0.81-0.866）
	1	0.877（0.849-0.909）	0.831（0.798-0.863）	0.823（0.793-0.853）

然而，在其余实验中，我们仅使用2次乳腺X线筛查作为输入。尽管使用3次乳房X线摄影筛查在表4中取得了良好结果，但这可能会增加输入视频中重复乳腺X线影像的数量，因为并非所有患者都接受过2次或更多次序贯筛查（如图1所示）。如第2.2节（数据预处理）所述，对于筛查次数不足的患者，我们复制其最近一次筛查的影像，以填补序列乳腺X线图像视频中的时间点空缺。由于数据集中接受过3次或更多次序贯筛查的患者数量较少（见图1），为避免通过复制乳腺X线影像来维持模型的准确性和性能，我们仅使用2次乳腺X线筛查。

表5比较了模型第1层中 SHIFT模块各种配置所获得的结果。我们研究了在 SHIFT模块中对查询/键卷积以及 alpha/beta 全连接层采用权重共享的效果。通过分析结果可以明显看出，权重共享虽然减少了可训练权重的数量，但降低了注意力模块的性能。这表明，为查询/键和 alpha/beta 分别设置可训练权重十分重要，因为它们负责从图像特征中捕获更多信息。此外，我们还按照原始论文 [21] 中的描述，对查询和值相加进行了实验。尽管结果颇具潜力，但我们的最终配置（如图3所示）仍然是最优设置。

表5. 第1层中不同配置下 SHIFT 的结果汇总

配置	1年AUC（95% CI)	2年AUC（95% CI)	> 2年AUC（95%置信区间）
SHIFT	0.887（0.86-0.917）	0.848（0.82-0.876）	0.844 (0.818-0.872)
权重共享–查询, Key	0.881（0.852-0.914）	0.835（0.803-0.865）	0.831（0.802-0.86）
权重共享 – Alpha, Beta	0.873 (0.844-0.903)	0.841（0.815-0.875）	0.838（0.808-0.866）
查询值增益	0.879（0.848-0.911）	0.841（0.81-0.874）	0.84（0.813-0.867）

表6汇总了深度学习架构中非局部网络的不同注意力机制以及我们的SHIFT模块的性能。结果表明，加入注意力机制能够提升风险预测模型的性能；非局部网络和SHIFT均符合这一结论，其中1年、2年和> 2年曲线下面积的结果均有所改善。

我们还在表6中研究了将注意力模块置于ResNet不同层的情况。为了从注意力机制中获得最大的效果和收益，我们假设较高分辨率的图像特征能够从注意力机制中获得更多益处，因为在较早阶段应用注意力机制，随后将使其作用扩展至较低分辨率的图像特征。结果验证了我们的假设：对于SHIFT，将注意力模块置于ResNet第1层的效果优于置于第2层的效果（对于非局部网络，当模块置于第1层时，由于计算量呈指数级增加，尤其是在筛查乳房X线摄影数量较多时，内存不足）。在所有列出的曲线下面积结果中，我们的SHIFT 方法均优于非局部注意力。

表6. 非局部网络、SHIFT 以及未应用注意力机制时的结果汇总。对于非局部网络和 SHIFT，相应的注意力模块分别被添加到两个不同阶段，即 Resnet-18 的第1层或第2层。

注意力机制 机制	ResNet 层	1年AUC（95% CI)	2年AUC（95% CI)	> 2年曲线下面积（95% CI)
-	-	0.879（0.85-0.91）	0.839（0.811-0.871）	0.835（0.81-0.866）
非局部	1		内存不足	
	2	0.881 (0.822-0.91)	0.840（0.811-0.875）	0.837（0.81-0.867）
SHIFT	1	0.887（0.86-0.917）	0.848（0.82-0.876）	0.844（0.818-0.872）
	2	0.880（0.85-0.911）	0.842（0.811-0.873）	0.839（0.812-0.867）

对于维度相同的输入，与非局部网络相比，SHIFT 在计算方面的效率显著更高。这得益于 SHIFT 利用加性注意力汇总图像信息。SHIFT 无需计算每对图像像素位置之间的相似度矩阵，从而大幅降低了内存和计算成本。为对此进行检验，我们计算了注意力模块插入同一 Resnet 层（第1层）时非局部网络和 SHIFT 的计算成本，结果列于表7。尽管全连接层中的参数数量略有增加，SHIFT 的效率仍显著更高，速度也更快。

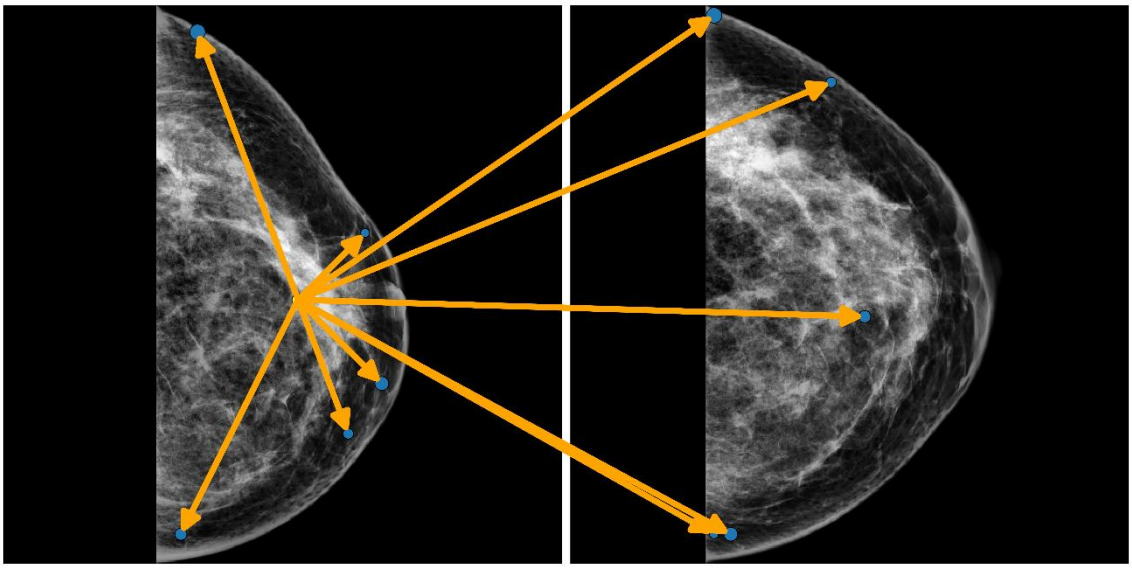
表7. 我们的三维（3D）网络架构中非局部模块与SHIFT模块的资源使用情况比较。

	非局部	SHIFT
计算量（乘加累积）：268.427M		67.682M
参数数量：8.192K		8.258K

乘加累积 = 乘加累积

图6展示了针对我们数据集中的一个癌症病例，分别由非局部块和我们的SHIFT模块计算得到的两张注意力图。这些注意力图直观地呈现了注意力机制关注当前图像（左侧）和先前1图像（右侧）的哪些部分。对于非局部块，使用致密组织区域附近的一个查询位置，并将加权值最高的10个点进行可视化。另一方面，SHIFT利用维度为P [illegible] 的注意力机制，为实

[illegible]



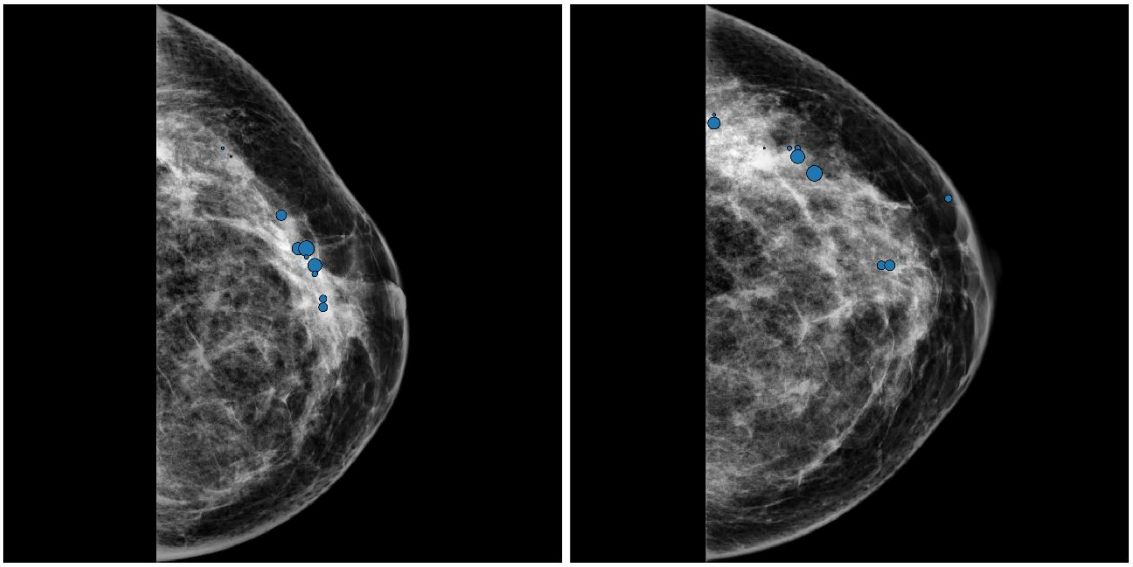


图6。（上）非局部注意力的注意力图。箭头表示在我们的数据集中，一名癌症患者的当前图像（左）和前1张图像（右）中，注意力从起始位置（查询位置）流向终点的方向。此处展示了加权点中权重最高的10个点。（下）同一名患者的SHIFT注意力图alpha。加权点按照权重从大到小的降序排列，并横跨两个乳腺X线摄影输入。

表8汇总了将影像组学特征纳入我们模型后的性能。我们在将影像组学特征输入全连接层之前，先将其与深度学习特征进行拼接。我们的测试表明，纳入基于频率的特征可以提升模型性能。在收集一组临床风险因素并不实际或不可行的情况下，这一点尤其有价值。在我们的具体案例中，利用影像组学特征对于增强模型的预测能力发挥了重要作用。

表8。将影像组学特征添加到我们模型中的结果汇总。影像组学特征在全连接层中与深度学习特征进行拼接。

影像组学特征		1年AUC（95%	2年AUC（95%	> 2-year AUC（95%
Pyradiomics	频率	CI)	CI)	CI)
-	-	0.879（0.85-0.91）	0.839（0.811-0.871）	0.835（0.81-0.866）
✓	-	0.875（0.842-0.91）	0.841（0.811-0.877）	0.841（0.814-0.87）
✓	✓	0.880（0.852-0.909）	0.845（0.817-0.876）	0.844（0.818-0.87）

除先前的研究结果外，我们还探索了一种名为 DAFT [31], 的替代特征组合方法，并将其应用于 ResNet架构的第三层。为评估其有效性，我们按照表9所示，比较了模型采用不同特征组合方法并结合 SHIFT 时的性能。尽管结果表明 DAFT 是一种可行的方法，但我们最终决定采用最初的特征组合方法，因为其总体性能最佳。

表9. 添加 SHIFT 后不同影像组学特征组合方法的结果汇总。

组合方法	SHIFT	1年曲线下面积（95% CI)	2年曲线下面积（95% CI)	> 2年曲线下面积（95% CI)
-	-	0.879（0.85-0.91）	0.839（0.811-0.871）	0.835（0.81-0.866）
全连接层	-	0.880 (0.852-0.909)	0.845 (0.817-0.876)	0.844（0.818-0.87）
	✓	0.888（0.86-0.919）	0.854（0.825-0.882）	0.853（0.827-0.877）

DAFT [31]	-	0.878（0.85-0.914）	0.858（0.832-0.888）	0.852（0.829-0.879）
	✓	0.886 (0.856-0.914)	0.857 (0.829-0.889)	0.853 (0.829-0.878)

接下来，表10展示了在模型中加入门控，使我们能够有效地对来自两个视图的信息进行加权。为评估其影响，我们尝试了不同的初始固定/可训练权重，以观察其效果。结果表明，无论选择何种初始权重，加入门控都能提升风险预测模型的性能，但部分权重设置的表现更佳。有趣的是，结果始终显示模型明显偏好头尾位视图，而非内外斜位视图；这与既往研究结果一致[9, 32], , 可能是因为内外斜位视图中存在胸大肌，从而对模型造成混淆。

表10. 门控机制不同初始权重的比较。

门控权重		1年曲线下面积（95%置信区间）	2年曲线下面积（95%置信区间）	> 2年曲线下面积（95% CI)
W_T, W_S	W_f			
0.5	0.5	0.888（0.858-0.918）	0.86（0.833-0.891）	0.858（0.832-0.886）
0.6	0.4	0.902（0.877-0.928）	0.863（0.835-0.891）	0.861（0.833-0.886）
0.7	0.3	0.894（0.868-0.925）	0.86 (0.83-0.889)	0.855（0.83-0.883）
0.8	0.2	0.886 (0.856-0.916)	0.863 (0.837-0.892)	0.859（0.832-0.885）

W_T = CC权重，W_S = MLO权重，W_f = 固定权重

利用我们的最佳模型，我们首先使用软标签重建完整数据集。接下来，我们应用双侧风险评分技术识别并移除离群对照组，从而得到优化后的筛选后数据集。如表11所示，仅使用其中任一数据集进行训练都会导致模型性能不佳。当模型在完整数据集上训练时，我们可以观察到2年AUC和>2年AUC有所提升，这可能归因于软标签。然而，最佳方法是在训练过程中同时利用两个数据集，从而获得更优的模型性能，表11中的结果也证实了这一点。

表11。我们的模型在不同版本数据集上的性能。最终迭代版本首先在完整数据集上训练，随后在筛选后数据集上进行微调。

数据集	1年AUC（95%置信区间）	2年AUC（95%置信区间）	> 2年曲线下面积（95%置信区间）
D_s	0.898（0.871-0.929）	0.867（0.839-0.897）	0.862 (0.837-0.888)
D_f	0.898 (0.871-0.925)	0.86 (0.832-0.889)	0.855 (0.829-0.882)
$D_s \rightarrow D_f$	0.905（0.88-0.931）	0.872 (0.846-0.898)	0.866（0.844-0.89）

D_s = 完整数据集，D_f = 筛选后数据集

表12列出了在模型中纳入新增改进后的结果。为将我们的方法与最先进的 Mirai 乳腺癌风险预测方法进行比较，我们根据其论文[8] 中提供的描述和超参数，以及 GitHub 上提供的代码，尽可能贴近地复现了作者在[8] 中的设置，并在我们的训练集上训练其模型，以进行公平比较。我们最终模型的性能更优，这主要可能归因于模型有效利用了患者的影像学历史（即此前的影像筛查）。作者也将此视为未来改进 Mirai 性能的一个方向/方法[8]。为从统计学上验证结果，我们进行了 DeLong检验[38], , 以我们的模型为参照，并将其与 Mirai 进行比较。检验结果显示，1年、2年和> 2年曲线下面积对应的 p 值分别为= 0.03、< 0.01 和<0.01，表明这两个模型之间存在显著差异。

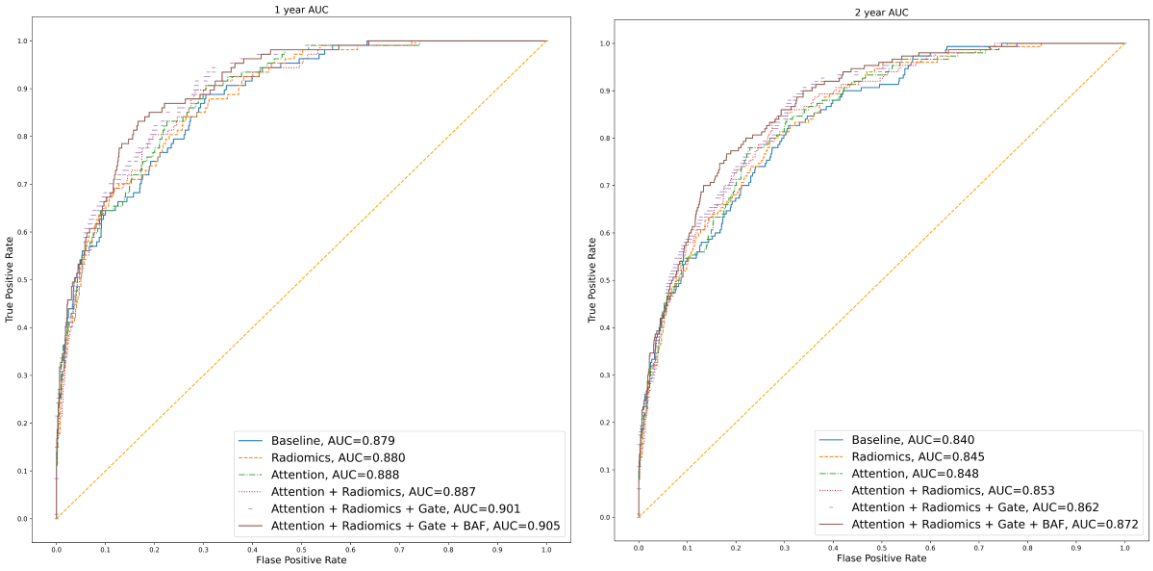
结果表明，结合深度学习、注意力机制、影像组学特征、门控和双侧微调的最终模型取得了最佳曲线下面积结果。结果还表明，SHIFT模块的自注意力机制能够捕捉不同序列筛查乳腺X线摄影图像中的有用信息，从而预测最近一次筛查乳腺X线摄影图像中的癌症风险；同时，深度学习和影像组学特征能够协同工作，提供

癌症风险预测的互补/补充信息。加入门控以重新加权每个乳腺影像视图的贡献，有利于模型性能。我们的结果还表明，吸收双侧不对称性的知识来微调模型，可以进一步提升性能。1年、2年和>2年曲线下面积结果的受试者工作特征（ROC）曲线绘制于图3中。

T表12。将新修改/新增内容纳入模型的结果汇总 .

模型配置	影像组学	1年AUC（95% CI)	2年AUC（95% CI)	> 2-年AUC (95%置信区间)
SHIFT + 门控 + BAF	✓	0.905 (0.88-0.931)	0.872 (0.846-0.898)	0.866 (0.844-0.89)
SHIFT + 门控	✓	0.902 (0.877-0.928)	0.863 (0.835-0.891)	0.861 (0.833-0.886)
SHIFT	✓	0.888 (0.86-0.919)	0.854 (0.825-0.882)	0.853 (0.827-0.877)
基线	✓	0.880 (0.852-0.909)	0.845 (0.817-0.876)	0.844 (0.818-0.87)
基线	-	0.879（0.85-0.91）	0.839 (0.811-0.871)	0.835 (0.81-0.866)
Mirai [8]	-	0.877（0.848-0.911）	0.834（0.804-0.865）	0.830（0.802-0.859）

SHIFT = 空间通道图像 Fastformer，BAF = 基于双侧不对称性的微调



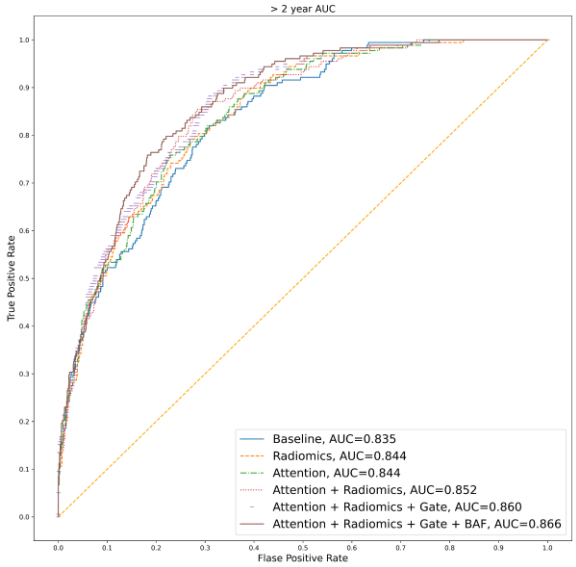


图7. 消融实验结果对应的1年、2年和>2年AUC类别的ROC曲线。

4. 讨论 在本文中，我们

报告了一项具有若干独特特点的新研究。首先，我们提出了一种深度学习模型，该模型能够利用注意力机制有效处理序列乳腺X线筛查，并将既往筛查纳入考量。序列乳腺X线图像可以显示乳腺组织随时间出现的异常或变化，而检测这些变化的能力对短期风险预测模型至关重要。纳入既往筛查也很关键，因为某些癌症比其他癌症更具侵袭性；越早发现癌症，患者的生存率就越高。

将多张乳腺X线影像组合形成视频输入后，我们采用三维自注意力方法和三维卷积神经网络处理这类输入。与传统二维卷积神经网络不同，三维卷积神经网络能够同时分析空间数据和时间数据，因此非常适合分析序列乳腺X线图像。通过仅扩展特定卷积核的方法（如第2.3节所述），可以将预训练现成二维卷积神经网络模型转换为三维卷积神经网络，同时仅增加少量参数。我们的结果表明，将自注意力用于检测不同筛查之间的全局信息，并结合三维卷积神经网络提取时空维度中的局部信息，这种方法能够优于仅使用单次筛查的先进风险模型。

我们的结果表明，筛查次数越多，模型表现越好，因为多次筛查能够提供更多信息；然而，使用更多次筛查会导致内存溢出问题，正如表6所示：在 ResNet 第1层引入非局部模块时，仅进行 2 次序贯筛查就已出现这一问题。因此，本研究的第二项创新贡献是提出一种新型基于图像的注意力机制，称为 SHIFT（空间通道图像 Fastformer）。我们此前针对其他器官/应用开展的研究表明，包括 [16] 中为视频分类引入的非局部注意力机制在内的注意力机制，均展现出良好效果 [22, 39]。注意力机制有助于对乳腺X线摄影图像中的不同区域赋予不同权重，从而预测乳腺癌风险。这在多次筛查的情况下尤为重要，注意力机制能够帮助捕捉乳腺X线影像之间的时空关系。认识到注意力机制的重要性后，我们着手开发一种能够处理高分辨率输入（即高分辨率序列乳腺X线图像）且取得较高曲线下面积（AUC）结果的注意力机制。

Fastformer [21] 最初是为自然语言处理开发的Transformer。与传统Transformer架构 [18], 中的自注意力相比，Fastformer具有线性复杂度，因此处理速度更快；而传统Transformer架构的自注意力具有二次复杂度和较高的内存开销。Fastformer从未应用于图像领域，我们的研究提出了一种新的快速图像注意力机制

。在我们的研究中，我们扩展了Fastformer原始论文的内容[21]，加入通道注意力（原始研究仅包含空间注意力）。我们还针对原始Fastformer架构开展了一些消融研究，包括查询添加；实验发现，查询添加是多余的。据我们所知，我们的研究首次测试了注意力机制在乳腺癌风险预测中的应用。结果表明，引入非局部或SHIFT均可改善用于风险预测的深度学习架构，其中SHIFT的曲线下面积表现更佳，运行速度也快得多。

第三，我们考察了利用离散余弦变换（DCT）和 FFT 变换生成的一些新型基于频率的影像组学特征，用于乳腺癌风险预测。这些新的统计特征能够捕捉与癌症风险相关的组织特征的额外信息，并且已被证明在良恶性分类中表现良好 [27]。结果表明，影像组学特征可补充用于风险预测的深度学习特征。通过结合影像组学特征和深度学习特征，可以捕捉乳腺组织纹理和模式的更多信息，从而提升性能。影像组学特征和深度学习特征可能捕捉组织的不同方面：影像组学提取组织纹理和乳腺密度信息 [14, 40], ，而深度学习能够学习传统影像组学特征难以提取的其他模式。我们的结果表明，结合这两类特征能够更全面地呈现组织模式和特征。以往研究往往单独使用其中一类特征 [14], ，但本研究结果表明，影像组学特征能够为深度学习特征提供补充信息，从而提升性能。

第四，我们的研究提出了一种新的轻量级门控机制，用于处理来自头尾位和内外斜位的深度学习特征。两种乳腺摄影视图都提供了与癌症风险预测相关的信息；然而，各视图的相对重要性可能有所不同，因为某些视图提供的信息多于其他视图 [9, 32, 35]。因此，结合两种视图的信息仍是最佳策略。结合多视图特征的一种常见方法是使用全连接层，文献中提出的不同策略均显示出良好效果 [32-34]。然而，全连接层中的参数数量可能会非常庞大，从而导致过拟合，并使模型的训练和评估在计算上开销巨大。我们提出的轻量级方法易于实现，与注意力机制类似，模型会学习不同的权重／缩放因子，以突出各视图的重要性，从而优化两种视图的组合并提升整体性能。与以往基于计算机算法的研究结果类似 [9, 32], ，我们的结果显示，在最终模型中，CC视图的权重高于MLO视图。

第五，借鉴我们先前研究的结果 [15, 29, 32], ，我们估算了左右乳房之间的双侧不对称性，并利用这些信息来改进/微调模型，以增强其能力。左右乳房在外观或密度上的不对称差异可能表明癌症发展，而机器学习模型可以通过训练来检测/量化这些差异，以进行风险预测。我们首次提出了一种改进训练数据集的方法，即移除双侧不对称性较高的对照组。为此，我们分别为每个独立视图生成风险评分，并据此计算不对称风险评分。随后，我们从训练集中筛除不对称评分较高的对照组，因为这些样本可能会影响/干扰网络训练。结果表明，先使用完整数据集上的软标签训练模型，再使用筛选后数据集上的软标签进行训练，可以提升模型性能。使用软标签进行训练有助于作为一种正则化技术，减少过拟合并提高模型的泛化能力。

第六，我们在一个全面的筛查数据集上检验了新方法，该数据集共包含 8,723 名患者，每位患者均有全部 4 个乳腺摄影视图（即 34,892 张乳腺X线影像），所得结果优于文献中最先进的方法。由于 CSAW 数据集包含多次序列筛查乳腺X线摄影图像，我们能够检验新模型预测短期癌症风险的表现；在精准医疗时代，这有助于为每位女性制定个性化乳腺筛查程序 [8, 15]。

尽管我们的研究结果令人鼓舞，也有新的发现，但我们承认，这是一项基于实验室的回顾性数据分析研究，存在一些可在未来研究中解决的局限性。首先，对于 CSAW 数据集中缺失的筛查乳腺 X 线摄影图像（因为患者未定期接受筛查），我们不得不进行复制，这可能影响了模型的表现。未来，我们计划在更大的数据集（或多个数据集）上评估模型表现，并确保每位患者都有规律、定期的筛查。其次，由于本研究属于开发性研究，我们没有评估其在临床测试中的效用，也没有评估其对辅助放射科医师的影响。第三，我们数据集中第 1 类的病例包括在筛查后 <= 60 天确诊癌症的患者；这些病例以 1 年 AUC 结果呈现[8]。若严格定义风险预测，这些病例或可剔除，但这样会大幅减少数据集中的病例数；我们可以考虑在未来使用更大的数据集对此进行研究。尽管存在这些局限性，我们的研究提出了一种新的深度学习架构，在全面的筛查数据集上结合了注意力机制、影像组学、多视图门控和基于双侧不对称性的方法，用于短期乳腺癌风险预测。

利益冲突声明

本文作者声明不存在任何利益冲突，也未与任何产品或服务可能与本文主题相关的公司存在任何关系。

致谢 本研究得到

基础研究基金计划（FRGS）和高等教育部的资助

马来西亚（MOHE），资助编号为 FRGS/1/2022/ICT02/MUSM/02/1。

数据可用性

作者未获准分享数据。

参考文献[1] A.N. Gia

quinto、H. Sung、K.D. Miller、J.L. Kramer、L.A. Newman、A.K. Minihan、A. Jemal、R.L. Siegel，《乳腺癌统计数据》，2022，CA：临床医师癌症杂志，72（2022）。[2] M.H. Gail、L.A. Brinton、D.P. Byar、D.K. Corle、S.B. Green、C. Schairer、J.J. Mulvihill，为每年接受检查的白人女性预测其个体乳腺癌发生概率，《美国国家癌症研究所杂志》，81 24（1989）1879-1886。[3] J.M. Cuzick，一种纳入家族及个人风险因素的乳腺癌预测模型，《临床实践中的遗传性癌症》，10（2004）A29 - A29。[4] A. Gastounioti、E.F. Conant、D. Kontos，超越乳腺密度：实质纹理分析在乳腺癌风险评估中作用日益增强的综述，《乳腺癌研究》，18（2016）91。[5] A. Gastounioti、S. Desai、V.S. Ahluwalia、E.F. Conant、D. Kontos，人工智能在乳腺癌风险乳腺X线表型分析中的应用：叙述性综述，《乳腺癌研究》，24（2022）14。[6] O.H. Maghsoudi、A. Gastounioti、C. Scott、L. Pantalone、F.-F. Wu、E.A. Cohen、S.J. Winham、E.F. Conant、C.M. Vachon、D. Kontos，Deep-LIBRA：一种用于稳健量化乳腺密度的人工智能方法，并在乳腺癌风险评估中进行独立验证，《医学图像分析》，73（2021）102138。[7] A. Yala、C. Lehman、T. Schuster、T. Portnoi、R. Barzilay，一种基于深度学习的乳腺摄影模型，可提升乳腺癌风险预测，《放射学》，292（2019）60-66。[8] A. Yala、P.G. Mikhael、F. Strand、G. Lin、K. Smith、Y. Wan、L. Lamb、K. Hughes、C. Lehman、R. Barzilay，迈向稳健的基于乳腺摄影的乳腺癌风险模型，《科学·转化医学》，13（2021）。[9] D. Arefan、A.A. Mohamed、W.A. Berg、M.L. Zuley、J.H. Sumkin、S. Wu，利用正常乳腺X线影像进行深度学习建模，以预测乳腺癌风险，《医学物理学》（2019）。[10] S. Dadsetan、D. Arefan、W.A. Berg、M.L. Zuley、J.H. Sumkin、S. Wu，利用纵向乳腺X线检查的深度学习进行乳腺癌风险预测，《模式识别》，132（2022）108919。[11] K. Loizidou、G. Skouroumouni、C. Nikolaou、C. Pitris，一种通过乳腺X线影像时间减影自动检测和分类乳腺微钙化的技术，IEEE Access，8（2020）52785-52795。[12] S. Timp、C. Varela、N. Karssemeijer，通过时间分析进行计算机辅助诊断，以改善放射科医师对乳腺X线肿块病变的解读，IEEE生物医学信息技术汇刊，14（2010）803-808。[13] T. Kooi、N. Karssemeijer，利用深度神经网络对对称差异和时间变化进行分类，以检测乳腺X线摄影中的恶性肿块，《医学影像学杂志》，4（2017）。[14] A. Anandarajah、Y. Chen、G.A. Colditz、A. Hardi、C.R.T. Stoll、S. Jiang，实质纹理研究与乳腺X线乳腺密度及乳腺癌风险的关联：对文献所用方法的系统综述，《乳腺癌研究》：BCR，24（2021）。[15] M. Tan、B. Zheng、J.K. Leader、D. Gur，乳腺X线图像特征变化与近期乳腺癌发生风险之间的关联，IEEE医学影像汇刊，35（2016）1719-1728。[16] X. Wang、R.B. Girshick、A. Gupta、K. He，非局部神经网络，2018 IEEE/CVF计算机视觉与模式识别会议，（2018）7794-7803。[17] A. Dosovitskiy、L. Beyer、A. Kolesnikov、D. Weissenborn、X. Zhai、T. Unterthiner、M. Dehghani、M. Minderer、G. Heigold、S. Gelly、J. Uszkoreit、N. Houlsby，一张图像胜过16×16个词：用于大规模图像识别的Transformer，arXiv，abs/2010.11929（2020）。[18] A. Vaswani、N.M. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A.N. Gomez、L. Kaiser、I. Polosukhin，《注意力就是一切》，arXiv，abs/1706.03762（2017）。[19] K. Dembrower、P. Lindholm、F. Strand，用于训练和评估深度神经网络的数百万级乳腺X线图像数据集及基于人群的筛查队列——筛查年龄女性队列（CSAW），《数字成像杂志》，33（2020）408-413。[20] J. Carreira、A. Zisserman，何去何从，动作识别？一种新模型与Kinetics数据集，2017 IEEE计算机视觉与模式识别会议（CVPR），（2017）4724-4733。

[21] C. Wu、F. Wu、T. Qi、Y. Huang，《Fastformer：加性注意力就是一切》，2021年。[22] M. Al-Shabi、K. Shak、M. Tan，《ProCAN：用于肺结节分类的渐进式增长通道注意力非局部网络》，《模式识别》，122（2020）108309。[23] A. Liew、C.C. Lee、B.L. Lan、M. Tan，CASPIANET++：《用于脑肿瘤分割的多维通道-空间非对称注意力网络与噪声学生课程学习范式》，《生物与医学中的计算机》，136（2021）104690。[24] J. Hu、L. Shen、S. Albanie、G. Sun、E. Wu，《挤压激励网络》，《IEEE模式分析与机器智能汇刊》，42（2017）2011-2023。[25] J. Fu、J. Liu、H. Tian、Z. Fang、H. Lu，《用于场景分割的双重注意力网络》，2019年IEEE/CVF计算机视觉与模式识别会议（CVPR），（2018）3141-3149。[26] H.J.W.L. Aerts、E.R. Velazquez、R.T.H. Leijenaar、C. Parmar、P. Grossmann、S. Carvalho、J. Bussink、R. Monshouwer、B. Haibe-Kains、D. Rietveld、F. Hoebers、M.M. Rietbergen、C.R. Leemans、A. Dekker、J. Quackenbush、R.J. Gillies、P. Lambin，《利用定量影像组学方法通过非侵入性成像解码肿瘤表型》，《自然·通讯》，5（2014）4006。[27] M. Heidari、S. Mirniaharikandehei、W. Liu、A.B. Hollingsworth、H. Liu、B. Zheng，《开发并评估一种预测恶性病例可能性的新型全局乳腺X线图像特征分析方案》，《IEEE医学影像汇刊》，39（2020）1235-1244。[28] M. Tan、J. Pu、B. Zheng，《使用序贯前向浮动选择（SFFS）和支持向量机（SVM）模型优化乳腺肿块分类》，《国际计算机辅助放射学与外科学杂志》，9（2014）1005-1020。[29] M. Tan、B. Zheng、P. Ramalingam、D. Gur，《基于双侧乳腺X线特征不对称性的近期乳腺癌风险预测》，《学术放射学》，20 12（2013）1542-1550。[30] J.J.M. van Griethuysen、A. Fedorov、C. Parmar、A. Hosny、N. Aucoin、V. Narayan、R. Beets-Tan、J.-C. Fillion-Robin、S.D. Pieper、H.J.W.L. Aerts，《用于解码放射影像表型的计算影像组学系统》，《癌症研究》，77 21（2017）e104-e107。[31] T. Wolf、S. Pölsterl、C. Wachinger、A.s.D.N. Initiative、A.I.B.L.f.s.o. ageing，《DAFT：在卷积神经网络中交织表格数据与三维图像的通用模块》，《神经影像》，260（2022）。[32] M. Tan、J. Pu、S. Cheng、H. Liu、B. Zheng，《用于预测近期乳腺癌风险的基于四视图乳腺X线图像特征的融合模型评估》，《生物医学工程年鉴》，43（2015）2416-2428。[33] M. Tan、J. Pu、B. Zheng，《使用计算机化乳腺X线图像特征分析方案减少假阳性召回》，《医学与生物学物理》，59（2014）4357 - 4373。[34] N. Wu、J. Phang、J. Park、Y. Shen、Z. Huang、M. Zorin、S. Jastrzebski、T. Févry、J. Katsnelson、E. Kim、S. Wolfson、U. Parikh、S. Gaddam、L. Lin、K. Ho、J.D. Weinstein、B. Reig、Y. Gao、H. Toth、K. Pysarenko、A. Lewin、J. Lee、K. Airola、E. Mema、S. Chung、E. Hwang、N. Samreen、S. Kim、L. Heacock、L. Moy、K. Cho、K.J. Geras，《深度神经网络提升放射科医师在乳腺癌筛查中的表现》，《IEEE医学影像汇刊》，39（2020）1184 - 1194。[35] A.A. Mohamed、Y. Luo、H. Peng、R.C. Jankowitz、S. Wu，《理解临床乳腺X线乳腺密度评估：深度学习视角》，《数字成像杂志》，31（2017）387-392。[36] D. Scutt、J.T. Manning、G.H. Whitehouse、S.J. Leinster、C.P. Massey，《乳房不对称性、乳房大小与乳腺癌发生之间的关系》，《英国放射学杂志》，70（1997）1017-1021。[37] D.P. Kingma、J. Ba，《Adam：一种随机优化方法》，CoRR，abs/1412.6980（2015）。[38] E.R. DeLong、D.M. DeLong、D.L. Clarke-Pearson，《比较两个或多个相关受试者工作特征曲线的曲线下面积：一种非参数方法》，《生物统计学》，44 3（1988）837-845。[39] M. Al-Shabi、B.L. Lan、W.Y. Chan、K.H. Ng、M. Tan，《使用深度局部-全局网络进行肺结节分类》，《国际计算机辅助放射学与外科学杂志》，14（2019）1815-1819。[40] M. Tan、S. Mariapun、C.H. Yip、K.H. Ng、S.-H. Teo，《一种利用亚洲队列乳腺X线图像实质纹理分析确定乳腺癌风险的新方法》，《医学与生物学物理》，64（2019）。

Yeoh Hong Hui 是马来西亚莫纳什大学工程学院的博士生。目前，他从事医学影像领域的深度学习研究，主要根据乳腺X线摄影图像预测乳腺癌风险。

Andrea Liew 是马来西亚莫纳什大学工程学院的博士研究生。她的研究兴趣包括使用深度学习技术进行脑部和乳腺病灶检测。

Raphaël Phan 是莫纳什大学教授，专长于机器学习、安全、密码学和恶意人工智能。他已发表超过 90 篇期刊论文和 120 多篇会议论文。

Fredrik Strand 是卡罗林斯卡大学医院乳腺影像科的副教授兼医学博士放射科医师。他的研究兴趣包括将新型机器学习技术应用于乳腺放射影像。

Kartini Rahmat 是马来亚大学医学中心的教授兼顾问放射科医师。她目前是多个乳腺肿瘤学和癌症影像组学拨款项目的首席研究员及拨款持有人。

Tuong Linh Nguyen 是人口与全球健康学院流行病学与生物统计学中心的研究员。他目前的研究领域是乳腺癌。

John L. Hopper 是遗传流行病学家及墨尔本大学教授，同时担任教授级研究员，并出任人口与全球健康学院流行病学与生物统计学中心主任。

Maxine Tan 是莫纳什大学工程学院的高级讲师。她的研究兴趣包括利用深度学习和定量影像生物标志物改进癌症筛查、诊断和预后评估。