arXiv:2304.00433v1 [eess.SP] 2 Apr 2023

利用生成对抗网络进行马尔可夫链蒙特卡罗理想观察者计算

Weimin Zhou、Umberto Villa 和 Mark A. Anastasio

摘要—医学成像系统通常通过客观或特定任务的图像质量（IQ）指标进行评估和优化，这些指标用于量化观察者在特定临床相关任务上的表现。贝叶斯理想观察者（IO）的性能为所有观察者（无论是数值观察者还是人类观察者）设定了上限，并被提议用作评价指标（FOM），以评估和优化医学成像系统。然而，IO检验统计量对应于似然比，在大多数情况下难以计算。此前已有一种采用马尔可夫链蒙特卡罗（MCMC）技术的基于采样的方法，用于估计 IO性能。然而，目前将 MCMC方法用于理想观察者近似的应用仅限于少数几种情形，在这些情形中，待成像对象的分布可由相对简单的随机对象模型（SOM）描述。因此，亟需扩展 MCMC方法的适用范围，以应对大量需要进行基于 IO 的评估、但尚无相应信号对象模型可用的场景。本研究介绍并评估了一种新颖的 MCMC 方法，该方法采用基于生成对抗网络（GAN）的 SOM，称为 MCMC-GAN。研究利用具有参考解的测试用例，对 MCMC-GAN 方法进行了定量验证。结果表明，MCMC-GAN 方法能够扩展 MCMC方法在医学成像系统 IO 分析中的适用范围。

索引词—贝叶斯理想观察者、马尔可夫链蒙特卡罗、生成对抗网络

I. 引言

人们普遍认为，应基于图像质量（IQ）的客观测量指标对医学成像系统进行评估和优化 [1]–[3]。IQ的客观测量指标量化了观察者执行特定任务的能力，这些任务可以是信号检测任务或参数估计任务 [2],[4]–[6]。作为计算信号检测任务图像质量客观测量指标的一种手段，人们积极探索了多种观察者检测任务 [7]–[26]。其中，贝叶斯理想观察者（理想观察者）之所以与众不同，是因为它采用完整的任务特定信息，并为所有观察者设定了性能上限 [2], [27], [28]。因此，IO性能被提倡用作评价指标（FOM），以评估成像系统 [1], [3]。借助这一评价指标，可以优化成像系统，使采集到的图像数据中包含的任务特定信息量最大化。IO性能还可用于计算其他次优数值观察者或人类观察者的效率 [29]。

在考虑二元信号检测任务时，IO 使用似然比来计算检验统计量 [2]。对于给定任务，二元信号检测任务中的 IO 可提供最佳的受试者工作特征（ROC）曲线，并使 ROC 曲线下面积（AUC）达到最大值[2], [30]。然而，除某些特殊情况外，IO 检验统计量无法用图像数据的线性函数表示，也无法通过解析方法确定。因此，关于成像系统 IO 分析的文献大多局限于背景精确已知（BKE）信号检测任务，以及涉及高斯背景的简单背景统计已知（BKS）信号检测任务 [27], [31]–[33]。

近年来，研究人员重新探讨了使用监督学习来近似 IO [34]，并取得了显著进展，成功用于处理更复杂的检测任务和检测-定位任务 [18], [35]。尽管这一研究方向前景可期，但相关研究仍在进行中；此外，某些检测估计任务 [25], 无法仅通过监督学习方法解决，因此需要采用其他统计抽样方法 [25]。

在 Kupinski 等人的一项开创性研究中，建立了一种采用马尔可夫链蒙特卡罗（MCMC）技术的基于采样的方法，用于在考虑某些随机对象模型（SOM）（例如团块对象随机对象模型 [36], ）时，对理想观察者检验统计量进行数值近似 [4]。SOM 是一种生成模型，可用于从预先设定的待成像对象统计分布中进行采样 [37], [38]。SOM 的典型用途是生成对象集合或其离散近似，以便通过虚拟成像研究计算图像质量的客观度量 [39], [40]。后来，这种 MCMC 方法经调整后，用于估计包含二值纹理模型 [31]和参数化躯干

本研究部分受NIH奖项EB031772（子项目6366）和EB031585资助。（通讯作者：Mark A. Anastasio和Weimin Zhou）

Weimin Zhou任职于中国上海上海交通大学全球未来技术研究院，电子邮箱：weimin.zhou@sjtu.edu.cn。

Umberto Villa任职于美国得克萨斯州奥斯汀78712号德克萨斯大学奥斯汀分校Oden计算工程与科学研究所，电子邮箱：uvilla@oden.utexas.edu。

Mark A. Anastasio 任职于伊利诺伊大学厄巴纳-香槟分校生物工程系，地址：美国伊利诺伊州厄巴纳，邮编 61801；电子邮箱：maa@illinois.edu。

体模 [41]的 SOM 的理想观察者检验统计量。该方法还被用于计算某些通道化理想观察者（CIOs） [13], [14]。然而，尚未研究适用于更通用 SOM 的基于 MCMC 的方法；此外，现有 SOM 仅涵盖部分医学成像模态和应用。这限制了目前使用 MCMC 方法对医学成像系统进行理想观察者分析。

深度生成模型（如生成对抗网络（GANs）） [42], 在建立用于描述对象有限维近似的信号对象模型（SOMs）方面具有巨大潜力。GAN由生成器和判别器组成，二者均由深度神经网络构成。生成器和判别器通过对抗过程进行联合训练。GAN的最新进展，如渐进式增长生成对抗网络（ProGANs） [43] 和基于风格的生成对抗网络 [44], [45],，使高分辨率图像的合成成为可能。当针对待成像对象的分布对GAN进行专门训练，使其能够从中采样时，生成器便可用作信号对象模型（SOM） [38],[46], [47]。

作为对Zhou和Anastasio近期一项研究的后续工作 [6], ，本文介绍并评估一种新颖的MCMC方法，该方法采用基于GAN的信号对象模型（SOM），称为 MCMC-GAN。MCMC-GAN方法旨在扩展用于理想观察者（IO）近似的MCMC方法的适用范围，使其能够应用于所需信号对象模型（SOM）尚不可用、但有望通过生成对抗网络（GAN）或其他深度生成模型建立的情形。本文使用具有参考解的测试用例对MCMC-GAN方法进行了定量验证。此外，还将该方法应用于一个没有参考解、且传统MCMC方法不适用的问题。

本文其余部分安排如下。第 II 节回顾二元信号检测理论的主要内容，以及用于计算 IO 的传统 MCMC 方法。第 III 节介绍所提出的 MCMC-GAN 方法。第 IV 节和第 V 节分别给出 MCMC-GAN 方法的数值研究结果，其中涉及块状对象模型和一组临床 MR 脑部图像。最后，本文在第 VI 节讨论所提出的 MCMC-GAN 方法的潜在优势以及未来研究主题。

II. 背景

考虑一项二元信号检测任务，要求观察者将成像系统记录的图像数据分类为满足无信号假设（H_0）或有信号假设（H_1）。在这些假设下，测量图像数据可描述为：

$$\begin{aligned} H_0 &: \mathbf{g} = \mathbf{H}\mathbf{f}_b + \mathbf{n}, \\ H_1 &: \mathbf{g} = \mathbf{H}(\mathbf{f}_b + \mathbf{f}_s) + \mathbf{n}, \end{aligned} \tag{1}$$

其中，$\mathbf{g} \in \mathbb{R}^M$ 表示由数字成像系统采集的测量图像数据，$\mathbf{n} \in \mathbb{R}^M$ 表示测量噪声，而$\mathbf{H}$ 表示将对象映射为图像数据的成像算子。$\mathbf{f}_b$ 和$\mathbf{f}_s$ 分别表示背景对象和待检测信号。为方便起见，背景对象和信号的成像版本分别记为$\mathbf{b} \equiv \mathbf{H}\mathbf{f}_b$ 和$\mathbf{s} \equiv \mathbf{H}\mathbf{f}_s$。

成像算子$\mathbf{H}$ 可描述不同的映射，具体取决于正向问题采用连续到离散（C-D）还是离散到离散（D-D）表述 [2]。当成像过程描述为 C-D 映射时（这是数字成像系统的恰当描述），$\mathbf{f}_b$ 和 $\mathbf{f}_s$ 是函数，而$\mathbf{H}$将函数映射为 M 维向量$\mathbf{g}$，该向量描述测量图像。如果将成像过程近似为 D-D 映射，则$\mathbf{f}_b$ 和$\mathbf{f}_s$ 的离散化近似表示为 N 维向量，而矩阵算子$\mathbf{H} \in \mathbb{R}^{M\times N}$ 将对象的有限维表示映射为$\mathbf{g}$。

对于信号精确已知/背景统计已知（SKE/BKS）检测任务，信号$\mathbf{f}_s$ 是确定性的，而背景对象$\mathbf{f}_b$ 是随机的，可用概率密度函数（PDF）$\mathrm{pr}(\mathbf{f}_b)$描述，该函数刻画了待成像受试者群体中的变异性。尽管通常无法直接获取此概率密度函数，但原则上可以建立一个 SOM，从中进行抽样 [37],[38]。因此，对于 SKE/BKS 检测任务，物体变异性和测量噪声是图像数据$\mathbf{g}$ 变异性的来源。本文将研究此类任务。

为执行二元信号检测任务，观察者计算检验统计量 $t(\mathbf{g})$，将测量图像数据$\mathbf{g}$ 映射为实数标量。将检验统计量 $t(\mathbf{g})$ 与预先确定的阈值 τ 进行比较，以判定$\mathbf{g}$ 属于两个假设中的哪一个。通过改变阈值 τ，可以绘制 ROC 曲线 [48] ，展示假阳性率（FPF）与真阳性率（TPF）之间的权衡。随后可计算 ROC曲线下面积（AUC），以量化观察者性能。

A. 贝叶斯理想观察者

贝叶斯理想观察者（IO）采用一种利用完整统计知识的决策策略，并为所有观察者设定了性能上限。对于二元信号检测任务，IO 所采用的检验统计量是似然比 $\Lambda(\mathbf{g})$ 的任意单调变换，其定义为 [2], [4], [34]

$$\Lambda(\mathbf{g}) = \frac{\mathrm{pr}(\mathbf{g}|H_1)}{\mathrm{pr}(\mathbf{g}|H_0)}, \tag{2}$$

其中，$\mathrm{pr}(\mathbf{g}|H_j)$ 是条件概率密度函数，用于描述退化后的$\mathbf{g}$ 在假设 H_j (j =0,1) 下的似然。对于大多数 BKS 任务，理想观察者检验统计量在解析上难以计算，因为条件概率密度函数 $\mathrm{pr}(\mathbf{g}|H_j)$ 的计算涉及 $\mathrm{pr}(\mathbf{f}_b)$，而后者并非显式已知。若有 SOM 可用于从 $\mathrm{pr}(\mathbf{f}_b)$ 中采样，则可采用 Kupinski等人 [4] 提出的马尔可夫链蒙特卡罗（MCMC）方法，近似计算理想观察者检验统计量，具体如下所述。

B. 马尔可夫链蒙特卡罗

考虑 aSKE 二元信号检测任务时，似然比 $\Lambda(\mathbf{g})$ 可计算为 [4]:

$$\Lambda(\mathbf{g}) = \frac{\int d\mathbf{b}\ \mathrm{pr}_b(\mathbf{b})\mathrm{pr}(\mathbf{g}|\mathbf{b}, H_1)}{\int d\mathbf{b}\ \mathrm{pr}_b(\mathbf{b})\mathrm{pr}(\mathbf{g}|\mathbf{b}, H_0)} \equiv \int d\mathbf{b}\ \Lambda_{\mathrm{BKE}}(\mathbf{g}|\mathbf{b})\mathrm{pr}(\mathbf{b}|\mathbf{g}, H_0), \quad (3)$$

其中，$\Lambda_{\mathrm{BKE}}(\mathbf{g}|\mathbf{b})$ 是给定背景图像数据 $\mathbf{b} = \mathbf{H}\mathbf{f}_b$ 时的似然比，而 $\mathrm{pr}(\mathbf{b}|\mathbf{g}\ H_0)$ 是后验概率密度函数。这些量可计算如下：

$$\Lambda_{\mathrm{BKE}}(\mathbf{g}|\mathbf{b}) = \frac{\mathrm{pr}(\mathbf{g}|\mathbf{b}, H_1)}{\mathrm{pr}(\mathbf{g}|\mathbf{b}, H_0)}, \quad (4a)$$

and

$$\mathrm{pr}(\mathbf{b}|\mathbf{g}, H_0) = \frac{\mathrm{pr}(\mathbf{g}|\mathbf{b}, H_0)\mathrm{pr}_b(\mathbf{b})}{\int d\mathbf{b}'\mathrm{pr}(\mathbf{g}|\mathbf{b}', H_0)\mathrm{pr}_b(\mathbf{b}')}, \quad (4b)$$

其中，$\mathrm{pr}_b(\mathbf{b})$ 是背景图像数据 $\mathbf{b}$ 的概率密度函数。考虑随机背景对象 $\mathbf{f}_b$ 可由一个 SOM 描述，该 SOM 由随机向量 $\boldsymbol{\theta}$ 表征，且该向量具有概率密度函数 $\mathrm{pr}(\boldsymbol{\theta})$。在这种情况下，可通过对 $\boldsymbol{\theta}$ 进行采样来生成背景图像数据 $\mathbf{b}$：即，$\mathbf{b} \equiv \mathbf{b}(\boldsymbol{\theta})$。根据这些量，随后可将式 (3) 中描述的似然比计算为 [4]:

$$\Lambda(\mathbf{g}) = \int d\boldsymbol{\theta}\ \Lambda_{\mathrm{BKE}}(\mathbf{g}|\mathbf{b}(\boldsymbol{\theta}))\mathrm{pr}(\boldsymbol{\theta}|\mathbf{g}, H_0). \quad (5)$$

可采用蒙特卡罗积分近似该积分，从而得到似然比 [4] 的估计值：

$$\Lambda(\mathbf{g}) \approx \frac{1}{J}\sum_{j=1}^{J}\Lambda_{\mathrm{BKE}}(\mathbf{g}|\mathbf{b}(\boldsymbol{\theta}^j)), \quad (6)$$

其中，样本 $\boldsymbol{\theta}^j$ 从后验概率函数 $\mathrm{pr}(\boldsymbol{\theta}|\mathbf{g}\ H_0)$ 中抽取，J 是用于近似该积分的样本数。为获得样本 $\boldsymbol{\theta}^j$，可使用 Metropolis-Hastings算法[49]生成一个平稳密度为 $\mathrm{pr}(\boldsymbol{\theta}|\mathbf{g}\ H_0)$ 的马尔可夫链。具体而言，选择一个初始向量 $\boldsymbol{\theta}^0$，并指定一个提议密度函数 $q(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^j)$。对于给定向量 $\boldsymbol{\theta}^j$，$j \geq 1$，从提议密度 $q(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^j)$ 中抽取候选样本 $\tilde{\boldsymbol{\theta}}$ 作为链中的下一个样本，并以概率 $\mathrm{Pr}_a(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^j, \mathbf{g})$ 接受，该概率定义如下：

$$\mathrm{Pr}_a(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^j, \mathbf{g}) = \min\left[1, \frac{\mathrm{pr}(\mathbf{g}|\mathbf{b}(\tilde{\boldsymbol{\theta}}), H_0)\mathrm{pr}(\tilde{\boldsymbol{\theta}})q(\boldsymbol{\theta}^j|\tilde{\boldsymbol{\theta}})}{\mathrm{pr}(\mathbf{g}|\mathbf{b}(\boldsymbol{\theta}^j), H_0)\mathrm{pr}(\boldsymbol{\theta}^j)q(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^j)}\right]. \quad (7)$$

如果候选向量 $\tilde{\boldsymbol{\theta}}$ 被接受，它就会被添加到马尔可夫链中：$\boldsymbol{\theta}^{j+1} = \tilde{\boldsymbol{\theta}}$；否则，$\boldsymbol{\theta}^{j+1} = \boldsymbol{\theta}^j$。

然而，目前 MCMC方法的应用仅限于较简单的信号对象模型，例如块状对象模型 [4], 、二值纹理模型 [31], 和参数化躯干体模 [41]。接下来，将介绍如何扩展该方法，使其适用于基于GAN的随机对象模型。

C. 生成对抗网络

生成对抗网络（GAN）已得到积极探索，并成功用于建立深度生成模型，以生成符合训练“真实”图像集合随机特性的新“虚假”图像图像 [42]。生成对抗网络由生成器和判别器组成，二者均由深度神经网络构成。生成器通过对抗过程与判别器对抗训练。训练完成后，生成器可用于生成能够体现“真实”图像统计特性的“虚假”图像。

当生成对抗网络在有限维背景对象集合$\mathbf{f}_b$上训练时，生成器会将随机潜向量$\mathbf{z} \in \mathbb{R}^k$ 映射为一个“伪造”的背景对象$\hat{}\mathbf{f}_b = G(\mathbf{z}; \boldsymbol{\Theta}_G)$。这里，$G(\cdot\ ; \boldsymbol{\Theta}_G): \mathbb{R}^k \to \mathbb{R}^N$ 是由一个带有权重向量 $\boldsymbol{\Theta}_G$的深度神经网络表示的映射函数；潜向量$\mathbf{z}$ 则从一种简单且已知的分布（例如正态分布）中采样。判别器由另一个带有权重向量 $\boldsymbol{\Theta}_D$ 的深度神经网络及映射函数 $D(\cdot\ ; \boldsymbol{\Theta}_D): \mathbb{R}^M \to \mathbb{R}$表示。判别器将图像映射为实值评分，用于区分“真实”和“伪造”图像。生成对抗网络通过生成器与判别器之间的双人极小极大博弈进行训练：

$$\min_{\boldsymbol{\Theta}_G}\max_{\boldsymbol{\Theta}_D} V(D, G) = E_{\mathbf{f}_b}[l\left(D(\mathbf{b}; \boldsymbol{\Theta}_D)\right)] + E_{\mathbf{z}}[l(1 - D\left(G(\mathbf{z}; \boldsymbol{\Theta}_G); \boldsymbol{\Theta}_D\right))], \quad (8)$$

其中，$l(\cdot)$ 是一个目标函数，其取决于具体的训练策略。当 $D(\cdot\ ; \boldsymbol{\Theta}_D)$ 和 $G(\cdot\ ; \boldsymbol{\Theta}_G)$ 具有足够的容量时，$P_{\hat{\mathbf{f}}_b} = P_{\mathbf{f}_b}$ 当极小极大博弈达到全局最优 [42]。这里，$P_{\mathbf{f}_b}$ 表示“真实”背景对象$\mathbf{f}_b$的分布，而$P_{\hat{\mathbf{f}}_b}$ 表示“伪造”背景对象$\hat{}\mathbf{f}_b$的分布。随后，生成器可以表示一个描述背景对象集合内部变异性的 SOM，并且可以计算由潜向量$\mathbf{z}$ 参数化的背景图像数据$\mathbf{b}$：

$$\hat{\mathbf{b}} = \mathbf{H}G(\mathbf{z}; \Theta_G) \equiv \hat{\mathbf{b}}(\mathbf{z}). \quad (9)$$

有时，也可以直接使用背景图像数据$\mathbf{b}$训练生成对抗网络。在这种情况下，生成器可以直接生成“伪造”的背景图像数据：

$$\hat{\mathbf{b}} : \hat{\mathbf{b}} = G(\mathbf{z}; \Theta_G) \equiv \hat{\mathbf{b}}(\mathbf{z}). \quad (10)$$

当考虑 CD 成像算子，使得无法直接应用生成对抗网络来建立对象模型时，这一点尤其有用。

III. 使用生成对抗网络对理想观察者进行马尔可夫链蒙特卡罗近似

此处，将第 II-B 节所述的 MCMC方法推广到使用式 9 或式 10 中定义的 GAN生成数据$\hat{}$，具体取决于所用的对象模型$\mathbf{b}$ 。考虑一个与式 3—5 类似的 SKE/BKS信号检测任务，退化图像数据$\hat{}\mathbf{g}$ 的理想观察者检验统计量可计算如下：

$$\Lambda(\hat{\mathbf{g}}) = \int d\mathbf{z}\Lambda_{\mathrm{BKE}}(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}))p(\mathbf{z}|\hat{\mathbf{g}}, H_0), \quad (11)$$

其中，$\Lambda_{\mathrm{BKE}}(\hat{}\mathbf{g}|\hat{}\mathbf{b}(\mathbf{z}))$ 和 $p(\mathbf{z}|\hat{}\mathbf{g}\ H_0)$ 可计算为：

$$\Lambda_{\mathrm{BKE}}(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z})) = \frac{p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}), H_1)}{p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}), H_0)}. \quad (12a)$$

$$p(\mathbf{z}|\hat{\mathbf{g}}, H_0) = \frac{p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}), H_0)p_z(\mathbf{z})}{\int d\mathbf{z}' p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}'), H_0)p_z(\mathbf{z}')}. \tag{12b}$$

随后可采用马尔可夫链蒙特卡罗模拟来近似计算似然比：

$$\Lambda(\hat{\mathbf{g}}) \approx \frac{1}{J}\sum_{j=1}^{J} \Lambda_{\mathrm{BKE}}(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}^j)), \tag{13}$$

其中，$\mathbf{z}^j$ 从后验分布 $p(\mathbf{z}|\hat{}\mathbf{g}\ H_0)$中采样。为构造一个从后验分布 $p(\mathbf{z}|\hat{}\mathbf{g}\ H_0)$中抽取样本的马尔可夫链，需要指定一个提议密度函数。给定当前样本 $\mathbf{z}^j$，通过从提议密度函数 $q(\tilde{}\mathbf{z}|\mathbf{z}^j)$中采样，提出一个候选潜向量 $\tilde{}\mathbf{z}$ ，并以如下接受概率将其纳入马尔可夫链：

$$p_{\mathrm{a}}(\tilde{\mathbf{z}}|\mathbf{z}^j, \hat{\mathbf{g}}) = \min\left[1, \frac{p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\tilde{\mathbf{z}}), H_0)p_z(\tilde{\mathbf{z}})q(\mathbf{z}^j|\tilde{\mathbf{z}})}{p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}^j), H_0)p_z(\mathbf{z}^j)q(\tilde{\mathbf{z}}|\mathbf{z}^j)}\right]. \tag{14}$$

此处，概率密度函数 $p_z(\cdot)$具有简单的解析形式，因为潜向量 $\mathbf{z}$ 是从已知分布（如正态分布）中采样的。

在本研究中，潜向量$\mathbf{z}$ 从正态分布中采样：$\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_k)$。提议密度函数 $q(\tilde{}\mathbf{z}|\mathbf{z}^j)$基于预条件 Crank–Nicolson（pCN）算法 [50]设计。具体而言，给定当前潜向量 $\mathbf{z}^j$，候选向量 $\tilde{}\mathbf{z}$ 按如下方式计算：

$$\tilde{\mathbf{z}} = \sqrt{1-\beta^2}\mathbf{z}^j + \beta\boldsymbol{\xi}, \tag{15}$$

其中，β 是步长，$\boldsymbol{\xi}$ 从正态分布 $\mathcal{N}(\mathbf{0}, \mathbf{I}_k)$中采样。鉴于式（15）中 pCN 提议的先验不变性，即

$$p_z(\tilde{\mathbf{z}})q(\mathbf{z}^j|\tilde{\mathbf{z}}) = p_z(\mathbf{z}^j)q(\tilde{\mathbf{z}}|\mathbf{z}^j), \tag{16}$$

式（14）中的接受概率可计算为

$$p_{\mathrm{a}}(\tilde{\mathbf{z}}|\mathbf{z}^j, \hat{\mathbf{g}}) = \min\left[1, \frac{p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\tilde{\mathbf{z}}), H_0)}{p(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}^j), H_0)}\right]. \tag{17}$$

IV. 数值研究

我们通过计算机模拟研究，考察所提出的 MCMC-GAN 方法对由生成对抗网络建立的 SOM 所对应的理想观察者检验统计量的近似能力。我们考虑了两个 SKE/BKS 二元信号检测任务，其对应的 SOM 具有不同的逼真程度。在第一项数值研究中，我们将所提出的 MCMC-GAN 方法应用于块状对象模型生成的对象。我们使用专为块状对象模型设计的传统马尔可夫链蒙特卡罗算法 [4] 验证了理想观察者性能。第二项数值研究将所提出的 MCMC-GAN 方法应用于一组临床脑部MR图像；现有的 SOM 无法描述这些图像，因此无法直接应用传统的 MCMC 方法。我们使用 ROC 曲线评估观察者性能。曲线拟合采用 Metz-ROC 软件 [51]，并使用“适当的”双正态模型 [52]，[53]。下文将详细介绍所考虑的两个信号检测任务。

A. 团块状背景下的信号检测任务

第一个 SKE/BKS 二元信号检测任务采用随机块状对象模型来模拟随机背景。所谓的团块状背景（LB）可描述为：

$$f_b(\mathbf{r}) = \sum_{n=1}^{N_b} l(\mathbf{r} - \mathbf{r}_n|a, w_b), \tag{18}$$

其中，N_b 是服从均值为 6 的泊松分布的团块随机数量，而 $l(\mathbf{r} - \mathbf{r}_n|a, w_b)$描述团块的形状，该形状由二维高斯函数建模：

$$l(\mathbf{r} - \mathbf{r}_n|a, w_b) = a\exp\left(-\frac{(\mathbf{r}-\mathbf{r}_n)^T(\mathbf{r}-\mathbf{r}_n)}{2w_b^2}\right). \tag{19}$$

这里，$a = 1$、$w_b = 8$和$\mathbf{r}_n$ 表示 n^{th} 团块的中心位置。团块位置是在 64×64的图像视野内从均匀分布中采样得到的。

信号由二维高斯函数建模：

$$f_s(\mathbf{r}) = a_s\exp\left(-\frac{(\mathbf{r}-\mathbf{r}_s)^T(\mathbf{r}-\mathbf{r}_s)}{2w_s^2}\right), \tag{20}$$

其中，$a_s = 0.3$ 为信号幅度，$w_s = 2.5$ 为信号宽度，$\mathbf{r}_s = [32, 32]^T$ 为信号位置，对应于视野中心。

团块模型用于表征连续物体。我们考虑了一个理想化的平行孔准直器成像系统，该系统可由线性 C-D 映射描述，用于生成退化图像。该成像系统可描述为与高斯点响应函数（PRF） [4]，[37]进行卷积：

$$h_m(\mathbf{r}) = \frac{\hbar}{2\pi w_h^2}\exp\left(-\frac{(\mathbf{r}-\mathbf{r}_m)^T(\mathbf{r}-\mathbf{r}_m)}{2w_h^2}\right), \tag{21}$$

其中，h_m（$\mathbf{r}$）是 PRF，用于描述退化图像中m^{th} 元素对位置$\mathbf{r}$处物体的灵敏度；$\hbar = 35$ 和 $w_h = 2$ 分别是 PRF 的高度和宽度。虚拟成像系统采集尺寸为 64×64的图像。背景图像数据 b_m 和信号图像数据s_m 的 m^{th} （$1 \le m \le 4096$）元素可计算如下：

$$b_m = \frac{a\hbar w_b^2}{w_h^2 + w_b^2}\sum_{n=1}^{N_b}\exp\left(-\frac{(\mathbf{r}_m-\mathbf{r}_n)^T(\mathbf{r}_m-\mathbf{r}_n)}{2(w_h^2+w_s^2)}\right), \tag{22}$$

and

$$s_m = \frac{a_s\hbar w_b^2}{w_h^2 + w_s^2}\exp\left(-\frac{(\mathbf{r}_m-\mathbf{r}_s)^T(\mathbf{r}_m-\mathbf{r}_s)}{2(w_h^2+w_b^2)}\right). \tag{23}$$

噪声采用均值为 0、标准差为 20 的独立同分布（i.i.d.）多元高斯分布建模。无噪声背景图像数据 $\mathbf{b}$、信号图像数据 $\mathbf{s}$以及含信号测量数据 $\mathbf{g}$ 的示例如图 1 所示。

采用渐进式增长训练策略的生成对抗网络（即 ProGAN）[43] 使用 10,000 个背景图像数据样本进行了训练。训练完成后，带有

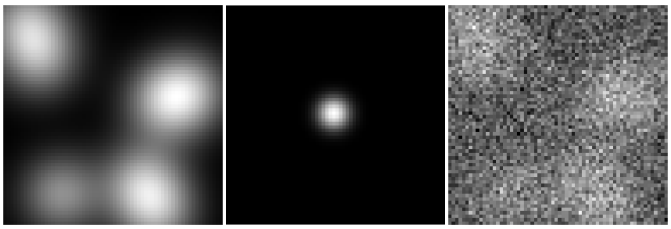


图 1：左：团块状背景图像数据。中：与所研究的信号检测任务对应的信号图像数据。右：含信号的噪声测量图像。

输入向量 $\mathbf{z}$ 可用于合成背景图像数据，并可根据公式 13 计算理想观察者（IO）检验统计量。由于考虑的是独立同分布的高斯噪声，因此 BKE似然比可计算为：

$$\Lambda_{\text{BKE}}(\hat{\mathbf{g}}|\hat{\mathbf{b}}(\mathbf{z}^j)) = \exp\left[(\hat{\mathbf{g}} - \hat{\mathbf{b}}(\mathbf{z}^j) - \mathbf{s}/2)^T K_n^{-1}\mathbf{s}\right], \quad (24)$$

其中， K_n 是对应于标准差为20的独立同分布高斯噪声的协方差矩阵。为计算理想观察者检验统计量，采用pCN算法构造马尔可夫链，其接受概率为 p_{a}（$\hat{}\mathbf{z}|\mathbf{z}^j$，$\mathbf{g}$），即高斯似然函数之比：

$$p_{\text{a}}(\tilde{\mathbf{z}}|\mathbf{z}^j, \mathbf{g}) = \min\left(1, \frac{\exp\left[-\frac{1}{2}(\hat{\mathbf{g}} - \hat{\mathbf{b}}(\tilde{\mathbf{z}}))^T K_n^{-1}(\hat{\mathbf{g}} - \hat{\mathbf{b}}(\tilde{\mathbf{z}}))\right]}{\exp\left[-\frac{1}{2}(\hat{\mathbf{g}} - \hat{\mathbf{b}}(\mathbf{z}^j))^T K_n^{-1}(\hat{\mathbf{g}} - \hat{\mathbf{b}}(\mathbf{z}^j))\right]}\right). \quad (25)$$

所提出的MCMC-GAN方法应用于200幅无信号图像和200幅有信号图像，以估计IO性能。为验证结果，采用针对团块状背景开发的传统MCMC方法（MCMC-LB） [4]作为理想观察者参考性能。作为进一步验证，采用协方差矩阵分解方法 [2]计算了所研究团块状背景下的Hotelling 观察者（HO）性能。

B. 使用临床*MRI*图像的信号检测任务

本研究考虑了针对临床脑部MR图像的信号已知统计量/背景已知统计量二元信号检测任务。采用了由阿尔茨海默病神经成像倡议（ADNI） [54] 资助的临床脑部MR数据集，作为真实背景对象$\mathbf{f}_b$集合。选取了12,000幅高质量矢状位脑部MR图像，并将其调整为 128×128的尺寸。随后，将这些图像归一化至0到1之间，用作训练ProGAN的训练图像。训练完成后，使用ProGAN的生成器通过对随机潜向量$\mathbf{z}$进行采样来合成背景对象。考虑了一个具有逼真脑肿瘤形状的信号对象，该对象选自脑肿瘤数据集（https:// figshare.com/articles/dataset/brain_tumor_dataset/1512427）。考虑了一个可获取欠采样k空间数据的风格化磁共振成像系统，并采用了变密度泊松盘采样模式 [55], [56]

其加速因子为16。此处考虑了对象的有限维近似，并将成像算子描述为D-D映射。通过将测量噪声添加到使用二维离散傅里叶变换（DFT）计算得到的欠采样k空间数据中，模拟测量数据$\mathbf{g}$ 。测量噪声$\mathbf{n}$ 建模为独立同分布、均值为零的复高斯随机向量，其实部和虚部的标准差均为 σ80。峰值信噪比（PSNR）定义为 $PSNR =$ 20 × 34 89 $\log_{10}\left(\frac{MAX_g}{\sigma}\right)$，其值为.dB。此处， MAX_g表示在测试数据集上测得的测量数据$\mathbf{g}$ 的最大值。图2展示了所考虑的MR脑图像$\mathbf{f}_b$、待检测信号$\mathbf{f}_s$以及k空间采样模式的示例。

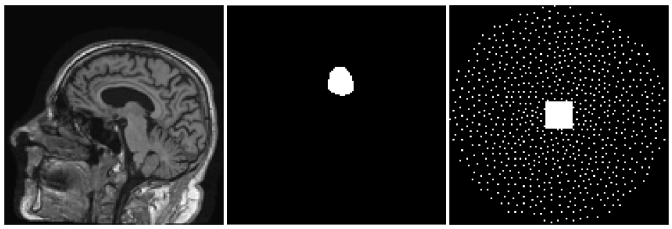


图 2：左：MR脑部对象的示例 $\mathbf{f}_b$。中：待检测的肿瘤信号。右：k空间采样模式。

所提出的 MCMC-GAN 方法应用于k空间数据，并在200组无信号测量数据和200组含信号测量数据上评估IO性能。采用pCN算法为每幅图像构建马尔可夫链。由于考虑的是独立同分布高斯噪声，因此分别根据公式24和25计算BKE似然比和提议接受概率。

由于所考虑的MR图像无法用现有的、可应用MCMC方法的SOM描述，因此未提供真实理想观察者性能。不过，我们实现了采用卷积神经网络（CNN）的监督学习方法 [16] ，以提供IO性能的参考值。训练CNN时，使用了包含一百万个由ProGAN生成的背景图像数据样本的训练数据集，以及一种“半在线学习”方法 [16] ，在该方法中，测量数据是在运行时生成的。我们将一个包含15个卷积（CONV）层的多通道卷积神经网络应用于零填充k空间数据的复值逆离散傅里叶变换，以估计IO性能。每个CONV层包含32个具有 5×5 空间支持的滤波器，并接有一个LeakyReLU激活函数。最后一个CONV层后接一个最大池化层和一个全连接（FC）层。此外，还通过采用协方差矩阵分解方法计算人类观察者（HO）性能，进行了额外验证。

C. MCMC-GAN 的收敛性分析

潜在尺度缩减因子（PSFR）是用于监测感兴趣的标量变量[57]–[59]的马尔可夫链蒙特卡罗收敛性的一项常用指标。PSFR基于并行链计算

，用于衡量序列内样本的平均方差与并行链中合并样本方差之比。设 $M > 1$ 表示链的数量，N_c 表示每条链中的样本数量，则PSFR可按以下方式计算 [57]–[59]:

$$\text{PSFR} = \sqrt{\frac{N_c - 1}{N_c} + \frac{1}{N_c}\frac{B}{W}}, \tag{26}$$

其中，W 为序列内方差，B 为序列间方差。这些量的计算如下:

$$W = \frac{1}{M(N_c - 1)} \sum_{m=1}^{M} \sum_{n=1}^{N_c} (v^{n,m} - \bar{v}^{\cdot,m})^2, \tag{27a}$$

$$\mathbf{B} = \frac{N_c}{M - 1} \sum_{m=1}^{M} (\bar{v}^{\cdot,m} - \bar{v}^{\cdots})^2, \tag{27b}$$

其中，$v^{n,\,m}$ 表示该 m^{th} 链的第 n^{th} 个样本，$\bar{v}^{\cdot,\,m}$ 表示来自 m^{th} 链的样本均值，$\bar{v}^{\cdot,\,\cdot}$ 。表示所有样本的均值:

$$\bar{v}^{\cdot,m} = \frac{1}{N_c} \sum_{n=1}^{N_c} v^{n,m}, \qquad \bar{v}^{\cdots} = \frac{1}{M} \sum_{m=1}^{M} \bar{v}^{\cdot,m}. \tag{28}$$

PSFR是一个实数标量，始终大于或等于1。当PSFR趋近于1时，马尔可夫链收敛到目标分布。文献 [59]建议采用1.01作为判断马尔可夫链是否收敛的阈值。

在本研究中，我们对BKE似然比 Λ_{BKE}计算了PSFR。BKE似然比是一个标量，用于计算蒙特卡罗积分，以近似求得理想观察者检验统计量。我们通过使用不同随机种子运行MCMC-GAN生成五条并行链，并据此计算PSFR。

D. MCMC-GAN 实现细节

ProGAN 使用 4 块 NVIDIA Quadro RTX 8000 GPU，通过 Tensorflow [60]进行训练。训练 ProGAN 时，采用了使用 Adam算法 [61] 的随机梯度法作为优化器。采用的 ProGAN 架构初始图像分辨率为 4×4，潜空间维度为 64。有关本研究所用 ProGAN 架构的更多细节，见 [43]。ProGAN 使用公开的 ProGAN代码（ https://github.com/tkarras/ progressive_growing_of _gans）进行训练。

ProGAN 训练完成后，采用带有 pCN 提议的 MCMC技术对 GAN 的潜变量进行采样，以计算理想观察者检验统计量。相应的提议密度函数和接受概率分别定义于式 15 和式 16。每条马尔可夫链均在单块 NVIDIA Quadro RTX 8000 GPU 上运行 200,000 次迭代。每条马尔可夫链的前 10,000 次迭代作为预热期予以舍弃。随后，根据式 13 对 190,000 次马尔可夫链迭代进行蒙特卡罗积分，进而计算与每份退化图像数据相对应的理想观察者检验统计量（即似然比）。对于每项所研究的信号检测任务，MCMC GAN 均应用于一组包含 200 张无信号图像和 200 张有信号图像的数据集，并使用所得检验统计量评估理想观察者的性能。

V. 结果

A. 团块状背景下的信号检测任务

真实值（上排）的团块状背景图像和 ProGAN 生成的图像（下排）如图 3 所示。ProGAN 生成的图像与真实值团块状背景图像的视觉外观相似。

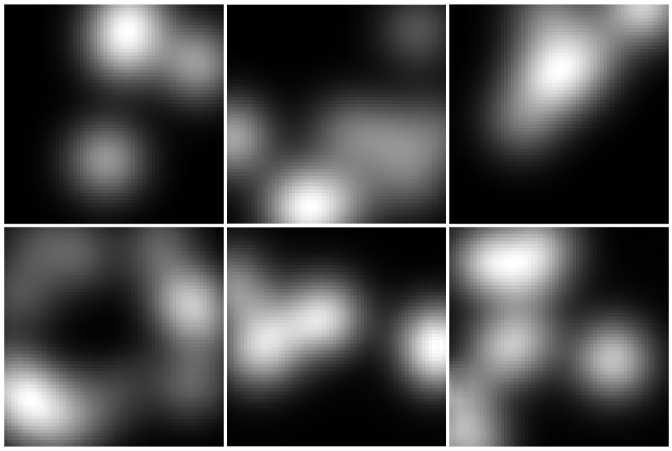


图 3：上：三个真实值团块状背景图像示例。下：三个 ProGAN 生成的图像示例。

用于评估 ProGAN 的功率谱 [62] 描述了“真实”图像和 ProGAN 生成的“伪”图像的频率成分。图 4 对比了 ProGAN 生成图像的功率谱（红色虚线）与“真实”图像的功率谱（蓝色曲线）。这两组功率谱分别对所有角度进行径向平均，并分别对 200 张“真实”和 200 张“伪”团块状背景图像取平均。两组功率谱几乎完全相同。

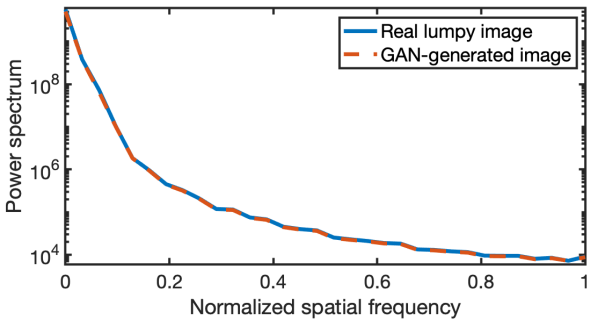
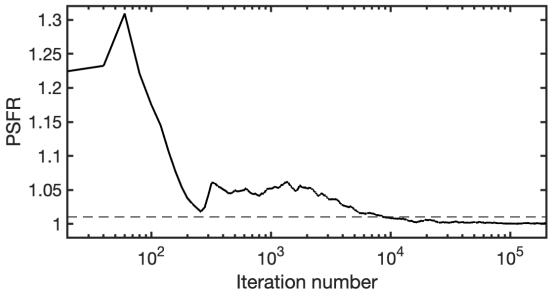


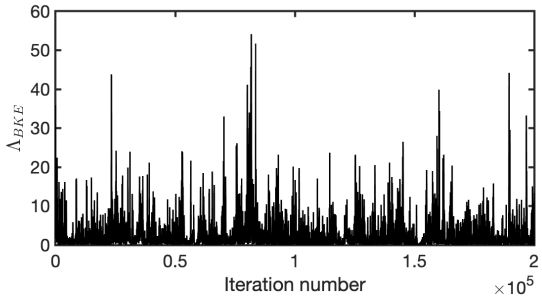
图 4： “真实”图像与 ProGAN 生成图像的功率谱。ProGAN 生成图像的功率谱与“真实”图像的几乎完全相同。

信号存在的团块图像中，Λ_{BKE} 样本的 PSFR 随马尔可夫链迭代次数变化的曲线

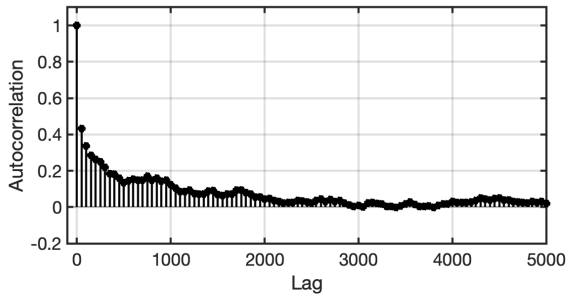
如图5(a)所示。图5(a)中的水平虚线表示收敛阈值 1.01。随着迭代次数增加，PSFR 趋近于 1，并在约 10,000 次迭代后收敛。链结束时的最终 PSFR 值为 1.0008。图5(b)展示了在马尔可夫链不同迭代次数下计算得到的 BKE似然比 Λ_{BKE} 链，其自相关函数绘于图5(c)中。



(a)



(b)



(c)

图5：（a）PSFR Λ_{BKE} 随马尔可夫链迭代次数变化的曲线。链 Λ_{BKE}约经过10,000次迭代后收敛。（b）在马尔可夫链不同迭代次数下评估的链 Λ_{BKE}。（c）（b）中链的自相关。

图6展示了 MCMC-GAN IO（蓝色曲线）、MCMC-LB IO（红色虚线曲线）和 HO（黄色曲线）对应的 ROC 曲线。MCMC-GAN IO、MCMC-LB IO 和 HO 对应的曲线下面积分别为 0.843 ± 0.019、0.840 ± 0.019 和 0.767 ± 0.023。MCMC-GAN IO 的性能与 MCMC-LB IO 非常接近，且优于 HO。

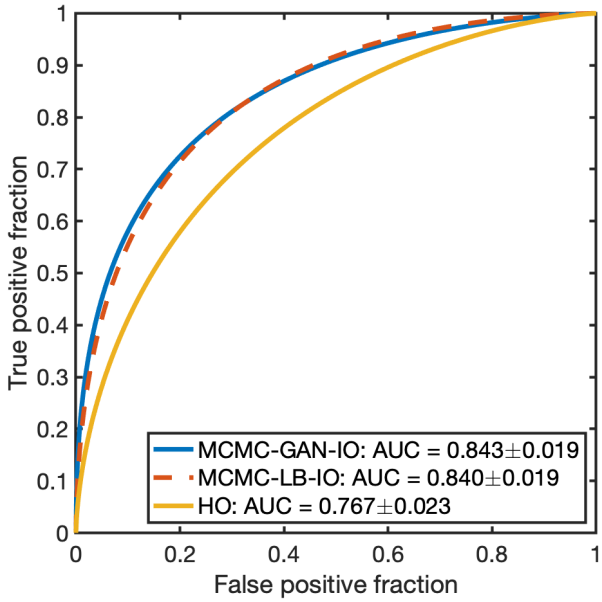


图6：MCMC-GAN-IO、MCMC-LB-IO和人类观察者生成的ROC曲线。MCMC-GAN-IO对应的ROC曲线与MCMC-LB-IO的曲线高度一致，且高于人类观察者的曲线。

B. 使用临床*MR*图像的信号检测任务

图7展示了真实值临床 MR脑图像（上排）和 ProGAN 生成的图像（下排）。ProGAN 生成的图像与真实值 MR图像的视觉外观相似。

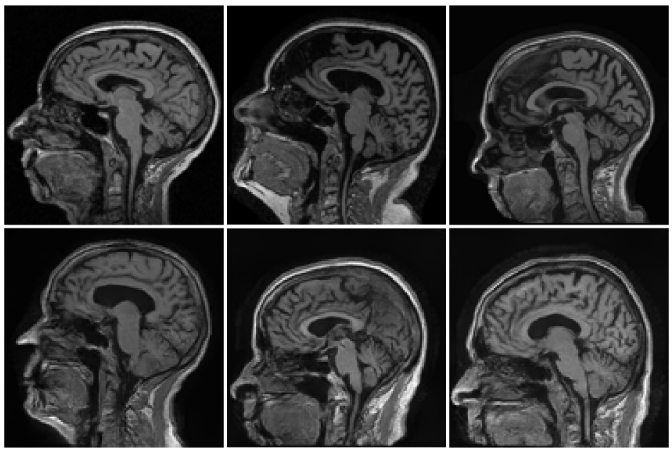


图7：上：真实值MR图像的三个示例。下：ProGAN生成图像的三个示例。

为评估ProGAN，我们进行了功率谱分析。图8展示了“真实”临床MR图像（蓝色曲线）和生成对抗网络生成的“伪造”MR图像（红色虚线）的功率谱。这些功率谱分别对所有角度进行了径向平均，并分别对200幅“真实”和200幅“伪造”MR脑部图像取平均。两个功率谱几乎完全相同。

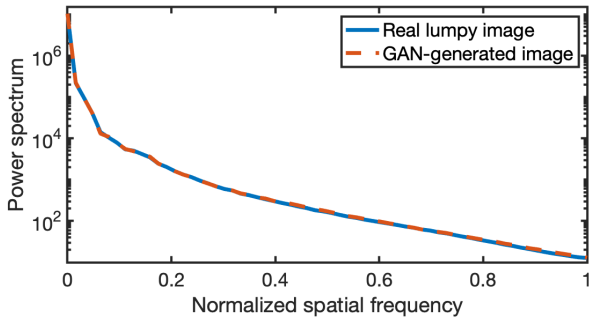


图8：真实图像和伪造图像的功率谱。这些功率谱几乎完全相同。

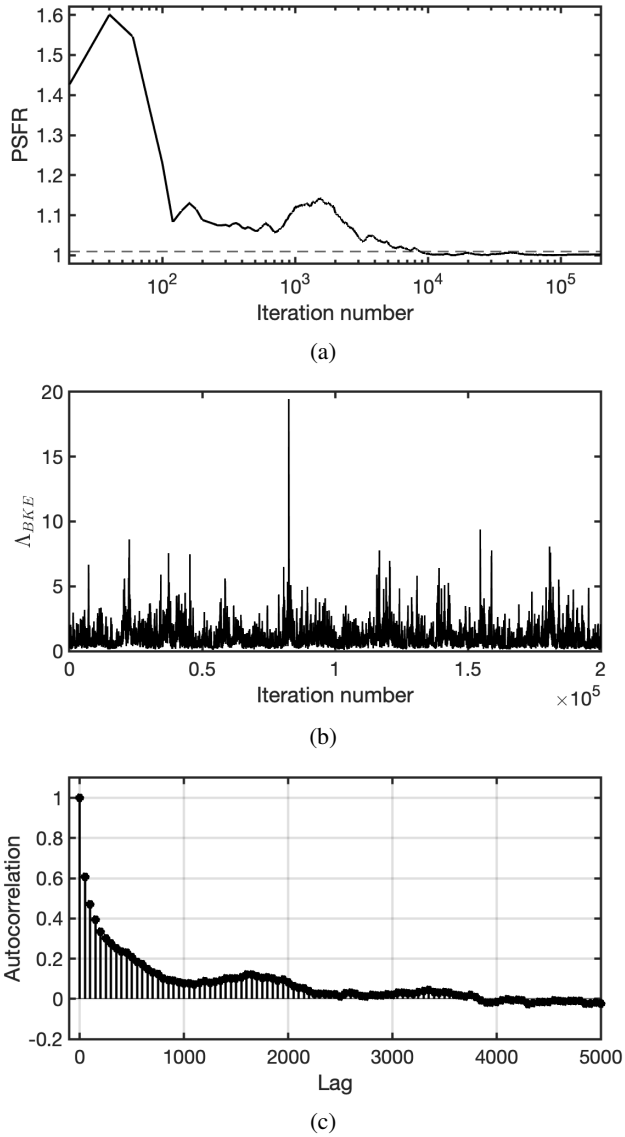


图9：(a) Λ_{BKE} 的 PSFR 随马尔可夫链迭代次数的变化曲线。该链在约 10,000 次迭代后 Λ_{BKE}收敛。(b) 在马尔可夫链不同迭代次数下评估的 Λ_{BKE}链。(c) (b) 中链的自相关。

图9(a)展示了信号存在的MR脑图像中， Λ_{BKE} 的PSFR随马尔可夫链迭代次数变化的曲线。图9(a)中的水平虚线表示1.01的收敛阈值。随着迭代次数增加，PSFR趋近于1，并在约10,000次迭代后收敛。链结束时的PSFR最终值为1.0026。图9(b)展示了在马尔可夫链不同迭代次数下计算的BKE似然比 Λ_{BKE} 。图9(c)绘出了该链的自相关。

图10展示了 MCMC-GAN IO（蓝色曲线）、CNN理想观察者（红色虚线）和人类观察者（黄色曲线）对应的 ROC曲线。MCMC-GAN IO、CNN理想观察者和人类观察者对应的曲线下面积分别为 0.866 ± 0.018、0.859 ± 0.018 和 0.821 ± 0.021。MCMC-GAN IO 的性能与 CNN理想观察者相近，且高于人类观察者。

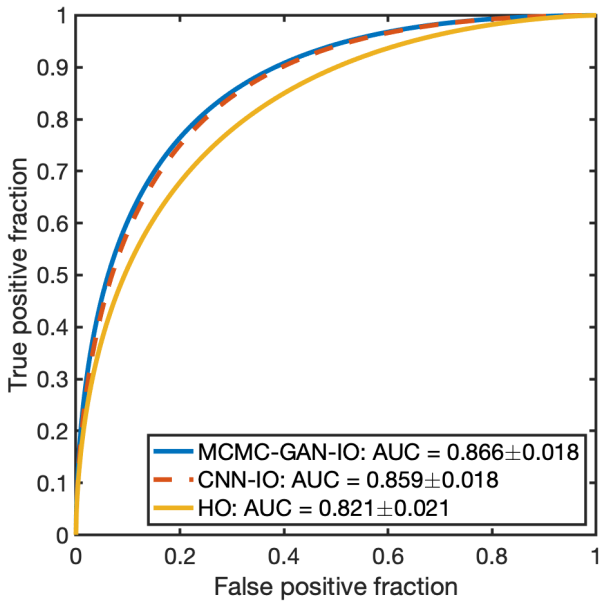


图 10：MCMC-GAN-IO、CNN-IO 和 HO 得到的 ROC 曲线。MCMC-GAN IO 对应的 ROC 曲线与 CNN-IO 的结果高度吻合，且高于 HO 的结果。

VI. 讨论与结论

利用完整任务特定信息的贝叶斯理想观察者（IO）已被广泛用于医学成像系统的评估与优化。然而，IO 的解析计算仅适用于一些特殊情况，而这些情况在实际应用中鲜少出现。为满足这一需求，Kupinski 等人提出了一种基于采样的方法，利用 MCMC 技术对块状对象模型的 IO 进行数值计算 [4]。该方法也已应用于其他一些 SOM，例如二值纹理模型 [31] 和参数化躯干体模[41]。然而，MCMC 方法应用于更

复杂的 SOM（能够表示真实医学图像中的物体变异性）的可行性仍缺乏充分研究。在本研究中，我们提出了一种新颖的基于采样的方法，即将 MCMC 技术与 GAN 建立的 SOM 结合（MCMC-GAN），以近似计算 IO。由于所提出的 MCMC-GAN 方法有望适用于任何由 GAN 建立的 SOM，因此 MCMC 技术在计算 IO 方面的适用范围得以扩展。为证明这一点，我们在一项数值研究中，将 MCMC-GAN 应用于一组临床脑部 MR 图像；这一研究无法通过传统 MCMC 方法完成。

本研究考察了二元信号检测任务，其中理想观察者检验统计量由似然比描述。所提出的 MCMC-GAN 方法也可用于联合信号检测与定位任务，在这类任务中，理想观察者由修正广义似然比检验描述 [63]。此外，MCMC-GAN 方法还可用于计算其他能够通过蒙特卡罗积分描述的量。MCMC-GAN 方法的另一种潜在应用是近似联合信号检测与估计任务中的理想观察者；在这类任务中，需要使用基于 MCMC 的方法计算一种称为效用加权后验均值的量 [24]。未来还需要探索 MCMC-GAN 近似联合信号检测与估计任务中理想观察者的能力。

未来仍有其他议题值得探究。在本研究中，马尔可夫链的样本是通过 pCN提议获得的 [50]。更先进的马尔可夫链蒙特卡罗算法，例如 Metropolis 调整的朗之万算法（MALA）和哈密顿蒙特卡罗（HMC） [64]，可以轻松地在我们提出的 MCMC-GAN 框架中实现。这是可行的，因为 GAN 建立的信号对象模型是一个神经网络，其关于潜向量的梯度可以在 TensorFlow 等机器学习平台上轻松计算 [60]。未来还需要研究采用其他 MCMC 采样器的 MCMC-GAN 的性能。此外，也可以采用 StyleGAN 等其他先进的生成对抗网络方法 [44], [45], [65] 和变分自编码器 [66]来建立用于 MCMC-GAN 方法的信号对象模型。

所提出的 MCMC-GAN 方法的一个局限是，它需要使用能够从真实物体分布中准确采样的生成对抗网络。然而，目前尚无全面评估生成对抗网络的方法，而在医学成像领域制定此类评估程序仍是一个持续研究的课题 [67]。因此，生成对抗网络能否可靠地捕捉理想观察者所需的物体统计特性仍不明确，这些特性因任务而异。不过，即使没有经过验证的生成模型，所提出的 MCMC-GAN 方法仍可用于信号检测任务，但可能只能近似理想观察者。

参考文献

[1] C. E. Metz、R. F. Wagner、K. Doi、D. G. Brown、R. M. Nishikawa 和 K. J. Myers，《医学成像系统定量评估的共识探索》，医学物理学，第 22 卷，第 7 期，第 1057–1061 页，1995 年。[2] H. H. Barrett 和 K. J. Myers，《图像科学基础》。John Wiley & Sons，2013 年。

[3] W. Vennart，*ICRU* 报告 *54*：医学成像中的图像质量评估：*ISBN 0-913394-53-X*。*1996*年*4*月，马里兰州，美国。WB Saunders，1997年。[4] M. A. Kupinski、J. W. Hoppin、E. Clarkson 和 H. H. Barrett，“使用马尔可夫链蒙特卡罗技术计算医学成像中的理想观察者”，*JOSA A*，第20卷，第3期，第430–438页，2003年。[5] F. Shen 和 E. Clarkson，“使用费 Fisher 信息近似理想观察者在块状背景图像检测任务中的性能”，*JOSA A*，第23卷，第10期，第2406–2414页，2006年。[6] W. Zhou 和 M. A. Anastasio，“使用生成对抗网络对理想观察者进行马尔可夫链蒙特卡罗近似”，载于*2020*年医学成像：图像感知、观察者性能与技术评估，第11316卷。国际光学与光子学学会，2020年，第113160D页。[7] H. H. Barrett、K. J. Myers、C. Hoeschen、M. A. Kupinski 和 M. P. Little，“基于任务的图像质量度量及其与辐射剂量和患者风险的关系”，《医学与生物学物理学》，第60卷，第2期，第R1页，2015年。[8] I. Reiser 和 R. Nishikawa，“基于任务的乳腺断层合成评估：采集参数与量子噪声的影响”，医学物理学，第37卷，第4期，第1591–1600页，2010年。[9] A. A. Sanchez、E. Y. Sidky 和 X. Pan，“通过 Hotelling 观察者度量优化专用乳腺 CT”，医学物理学，第41卷，第10期，2014年。[10] S. J. Glick、S. Vedantham 和 A. Karellas，“使用平板成像器研究 CT 乳腺摄影的最佳 kVp 设置”，载于医学成像 *2002*：医学成像物理学，第4682卷。国际光学与光子学学会，2002年，第392–403页。[11] H. H. Barrett、C. K. Abbey、B. D. Gallas 和 M. P. Eckstein，“患者结构化噪声中 Hotelling 观察者检测性能的稳定估计”，载于医学成像 *1998*：图像感知，第3340卷。国际光学与光子学学会，1998年，第27–44页。[12] B. D. Gallas 和 H. H. Barrett，“验证使用通道估计理想线性观察者的方法”，*JOSA A*，第20卷，第9期，第1725–1738页，2003年。[13] S. Park、H. H. Barrett、E. Clarkson、M. A. Kupinski 和 K. J. Myers，“在涉及非高斯分布块状背景和高斯信号的检测任务中使用拉盖尔-高斯通道的通道化理想观察者”，*JOSA A*，第24卷，第12期，第B136–B150页，2007年。[14] S. Park 和 E. Clarkson，“用于非高斯块状背景检测任务中通道化理想观察者性能评估的马尔可夫链蒙特卡罗”，载于医学成像 *2008*：图像感知、观察者性能与技术评估，第6917卷。国际光学工程学会，2008年，第236–243页。[15] ——，“高效估计涉及高维复杂背景的分类任务中的理想观察者性能”，*JOSA A*，第26卷，第11期，第B59–B71页，2009年。[16] W. Zhou、H. Li 和 M. A. Anastasio，“使用监督学习方法近似二元信号检测任务中的理想观察者和 Hotelling 观察者”，*IEEE*医学影像汇刊，第38卷，第10期，第2456–2468页，2019年。[17] W. Zhou 和 M. A. Anastasio，“使用卷积神经网络学习 SKE 检测任务中的理想观察者”，载于医学成像 *2018*：图像感知、观察者性能与技术评估，第10577卷。国际光学与光子学学会，2018年，第1057719页。[18] W. Zhou、H. Li 和 M. A. Anastasio，“使用监督学习方法近似联合信号检测与定位任务中的理想观察者”，*IEEE*医学影像汇刊，第1–1页，2020年。[19] ——，“使用监督学习方法学习 SKE 检测任务中的 Hotelling 观察者”，载于*2019*年医学影像：图像感知、观察者性能与技术评估，第10952卷。国际光学与光子学学会，2019年，第1095208页。[20] S. He、W. Zhou、H. Li 和 M. A. Anastasio，“使用无监督领域自适应学习数值观察者”，载于*2020*年医学成像：图像感知、观察者性能与技术评估，第11316卷。国际光学与光子学学会，2020年，第113160W页。[21] J. L. Granstedt、W. Zhou 和 M. A. Anastasio，“使用双损失自编码器学习高效通道”，载于*2020*年医学成像：图像感知、观察者性能与技术评估，第11316卷。国际光学与光子学学会，2020年，第113160C页。[22] E. Y. Sidky、J. P. Phillips、W. Zhou、G. Ongie、J. Cruz-Bastida、I. S. Reiser、M. A. Anastasio 和 X. Pan，“一种用于量化非线性图像重建中过度正则化的信号检测模型

。arXiv 预印本 arXiv:2010.03472，2020年。[23] J. P. Phillips、E. Y. Sidky、G. Ongie、W. Zhou、J. Cruz-Bastida、I. S. Reiser、M. A. Anastasio 和 X. Pan，“一种用于估计基于总变分的图像重建中理想线性观察者的混合通道化霍特林观察者”，载于《2021年医学影像：图像感知、观察者性能与技术评估》，第11599卷。国际光学与光子学学会，2021年，第115990D页。[24] K. Li、W. Zhou、H. Li 和 M. A. Anastasio，“用于联合检测与估计任务的基于监督学习的理想观察者近似”，载于《2021年医学影像：图像感知、观察者性能与技术评估》，第11599卷。国际光学与光子学学会，2021年，第115990F页。[25] ——，“一种结合监督学习与马尔可夫链蒙特卡罗方法近似联合信号检测与估计任务中理想观察者的混合方法”，IEEE 医学影像汇刊，第41卷，第5期，第1114–1124页，2021年。[26] ——，“评估基于深度神经网络的图像去噪对二元信号检测任务的影响”，IEEE 医学影像汇刊，第40卷，第9期，第2295–2305页，2021年。[27] K. J. Myers，“视觉信号检测的理想观察者模型”，《医学成像手册》，第1卷，第559–592页，2000年。[28] W. S. Geisler，“理想观察者分析”，《视觉神经科学》，第10卷，第7期，第12–12页，2003年。[29] A. Burgess、R. Wagner、R. Jennings 和 H. B. Barlow，“人类视觉信号辨别的效率”，《科学》，第214卷，第4516期，第93–94页，1981年。[30] R. F. Wagner 和 D. G. Brown，“医学成像系统的统一信噪比分析”，《医学与生物学物理学》，第30卷，第6期，第489页，1985年。[31] C. K. Abbey 和 J. M. Boone，“乳腺实质组织中 X射线成像模型的理想观察者”，载于《国际数字乳腺摄影研讨会》。施普林格，2008年，第393–400页。[32] A. K. Jha、E. Clarkson 和 M. A. Kupinski，“用于研究漫射光学成像中信号可检测性的理想观察者框架”，《生物医学光学快报》，第4卷，第10期，第2107–2123页，2013年。[33] M. A. Anastasio、C.-Y. Chou、A. M. Zysk 和 J. G. Brankov，“采用线性移不变光学系统的相位衬度成像中理想观察者信号可检测性的分析”，JOSA A，第27卷，第12期，第2648–2659页，2010年。[34] M. A. Kupinski、D. C. Edwards、M. L. Giger 和 C. E. Metz，“使用贝叶斯分类神经网络近似理想观察者”，IEEE医学影像汇刊，第20卷，第886–899页，第9期，2001年。[35] W. Zhou 和 M. A. Anastasio，“使用卷积神经网络学习联合检测与定位任务中的理想观察者”，载于《2019年医学影像：图像感知、观察者性能与技术评估》，第10952卷。国际光学与光子学学会，2019年，第1095209页。[36] J. P. Rolland 和 H. H. Barrett，“随机背景不均匀性对观察者检测性能的影响”，《美国光学学会杂志A辑》，第9卷，第5期，第649–658页，1992年5月。[在线]。可访问： http://josaa.osa.org/abstract.cfm?URI = 9 5 649 josaa- - - M. A.

[37] Kupinski、E. Clarkson、J. W. Hoppin、L. Chen 和 H. H. Barrett，“从含噪图像中实验确定对象统计特性”，JOSA A，第20卷，第3期，第421–429页，2003年。[38] W. Zhou、S. Bhadra、F. J. Brooks、H. Li 和 M. A. Anastasio，“使用先进的环境生成对抗网络从医学成像测量中学习随机对象模型”，《医学成像杂志》，第9卷，第1期，第015503页，2022年。[39] A. Badano，“计算机模拟成像临床试验：成本更低、速度更快、效果更好、更安全且更具可扩展性”，《临床试验》，第22卷，第1期，第1–7页，2021年。[40] E. Samei、E. Abadi、A. Kapadia、J. Lo、M. Mazurowski 和 P. Segars，“虚拟成像试验：成像研究与实践中一种新兴的实验范式”，载于《2020年医学影像：医学成像物理学》，第11312卷。国际光学工程学会，2020年，第439–444页。[41] X. He、B. S. Caffo 和 E. C. Frey，“面向医学成像系统优化的实用逼真理想观察者（IO）估计”，IEEE医学影像汇刊，第27卷，第10期，第1535–1543页，2008年。[42] I. Goodfellow、J. Pouget-Abadie、M. Mirza、B. Xu、D. Warde-Farley、S. Ozair、A. Courville 和 Y. Bengio，“生成对抗网络”，载于《神经信息处理系统进展》，2014年，第2672–2680页。[43] T. Karras、T. Aila、S. Laine 和 J. Lehtinen，“用于提升质量、稳定性和多样性的 GAN 渐进式增长”，arXiv 预印本 arXiv:1710.10196，2017。

[44] T. Karras、S. Laine 和 T. Aila，《用于生成对抗网络的基于风格的生成器架构》，载于《IEEE计算机视觉与模式识别会议论文集》，2019年，第4401–4410页。[45] T. Karras、S. Laine、M. Aittala、J. Hellsten、J. Lehtinen 和 T. Aila，《分析并提升 StyleGAN 的图像质量》，arXiv 预印本 arXiv:1912.04958，2019年。[46] W. Zhou、S. Bhadra、F. Brooks 和 M. A. Anastasio，《利用噪声成像测量学习随机对象模型：AmbientGAN》，载于《2019年医学影像：图像感知、观察者性能与技术评估》，第10952卷。国际光学与光子学学会，2019年，第109520M页。[47] W. Zhou、S. Bhadra、F. J. Brooks、J. L. Granstedt、H. Li 和 M. A. Anastasio，《推进 AmbientGAN 以学习随机对象模型》，载于《2021年医学影像：图像感知、观察者性能与技术评估》，第11599卷。国际光学与光子学学会，2021年，第115990A页。[48] C. E. Metz，《放射影像中的 ROC 方法》。Investigative Radiology，第21卷，第9期，第720–733页，1986年。[49] S. Chib 和 E. Greenberg，《理解 Metropolis-Hastings 算法》，The American Statistician，第49卷，第4期，第327–335页，1995年。[50] S. L. Cotter、G. O. Roberts、A. M. Stuart 和 D. White，《函数的马尔可夫链蒙特卡罗方法：改进旧算法以提高速度》，Statistical Science，第28卷，第3期，第424–446页，2013年。[51] C. Metz，《Rockit 用户指南》，芝加哥，放射学系，芝加哥大学，1998年。[52] C. E. Metz 和 X. Pan，《“正规”双正态 ROC 曲线：理论与最大似然估计》，Journal of Mathematical Psychology，第43卷，第1期，第1–33页，1999年。[53] L. L. Pesce 和 C. E. Metz，《可靠且计算高效的“正规”双正态 ROC 曲线最大似然估计》，Academic Radiology，第14卷，第7期，第814–829页，2007年。[54] S. G. Mueller、M. W. Weiner、L. J. Thal、R. C. Petersen、C. Jack、W. Jagust、J. Q. Trojanowski、A. W. Toga 和 L. Beckett，《阿尔茨海默病神经成像倡议》，Neuroimaging Clinics，第15卷，第4期，第869–877页，2005年。[55] M. Lustig、D. Donoho 和 J. M. Pauly，《稀疏 MRI：压缩感知在快速磁共振成像中的应用》，Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine，第58卷，第6期，第1182–1195页，2007年。[56] M. Uecker、F. Ong、J. I. Tamir、D. Bahri、P. Virtue、J. Y. Cheng、T. Zhang 和 M. Lustig，《伯克利高级重建工具箱》，载于《国际磁共振医学学会论文集》，第23卷，第2486期，2015年。[57] A. Gelman 和 D. B. Rubin，《利用多条序列进行迭代模拟推断》，Statistical Science，第7卷，第4期，第457–472页，1992年。[58] S. P. Brooks 和 A. Gelman，《监测迭代模拟收敛性的一般方法》，Journal of Computational and Graphical Statistics，第7卷，第4期，第434–455页，1998年。[59] A. Vehtari、A. Gelman、D. Simpson、B. Carpenter 和 P.-C. B¨urkner，《秩归一化、折叠与局部化：用于评估马尔可夫链蒙特卡罗收敛性的改进型 R^（含讨论）》，Bayesian Analysis，第16卷，第2期，第1–26页，2020年。[60] M. Abadi、P. Barham、J. Chen、Z. Chen、A. Davis、J. Dean、M. Devin、S. Ghemawat、G. Irving、M. Isard 等，《TensorFlow：一个大规模机器学习系统》，载于 OSDI，第16卷，2016年，第265–283页。[61] D. P. Kingma 和 J. Ba，《Adam：一种随机优化方法》，arXiv 预印本 arXiv:1412.6980，2014年。[62] F. O. Bochud、C. K. Abbey 和 M. P. Eckstein，《具有簇状团块背景的乳腺X线图像统计纹理合成》，Optics Express，第4卷，第1期，第33–43页，1999年。[63] P. Khurd 和 G. Gindi，《使 LROC 曲线下面积最大的决策策略》，IEEE医学影像汇刊，第24卷，第12期，第1626–1636页，2005年。[64] M. Pereyra、P. Schniter、E. Chouzenoux、J.-C. Pesquet、J.-Y. Tourneret、A. O. Hero 和 S. McLaughlin，《信号处理中随机模拟与优化方法综述》，IEEE信号处理专题期刊，第10卷，第2期，第224–241页，2015年。[65] T. Karras、M. Aittala、S. Laine、E. H¨ark¨onen、J. Hellsten、J. Lehtinen 和 T. Aila，《无混叠生成对抗网络》，《神经信息处理系统进展》，第34卷，第852–863页，2021年。[66] D. P. Kingma 和 M. Welling，《自编码变分贝叶斯》，arXiv 预印本 arXiv:1312.6114，2013年。[67] V. A. Kelkar、D. S. Gotsis、F. J. Brooks、K. Prabhat、K. J. Myers、R. Zeng 和 M. A. Anastasio，《评估生成对抗网络学习典型医学图像统计特性的能力》，IEEE医学影像汇刊，第1–1页，2023年。