噪声排除模型下二元贝叶斯分类器的似然比

Howard C. Gifford *

2026年1月14日

arXiv:2601.07982v1 [cs.CV] 12 Jan 2026

*The 作者现供职于美国德克萨斯州休斯顿的休斯顿大学。电子邮箱：hgifford@uh.edu

摘要

我们提出了一种新的统计理想观察者模型，通过对可提取的图像特征设定最低阈值，部分实现整体视觉搜索（或概要）处理。在该模型中，理想观察者减少了自由参数的数量，从而简化了系统。这一新框架可用于医学图像感知（以优化成像系统和算法）、计算机视觉、表现基准测试以及特征选择与评估。其他应用还包括国防/安全领域的目标检测与识别，以及传感器和探测器的评估。

1 引言

在医学成像技术发展的多个阶段，对诊断图像质量进行评估都能提供有价值的信息，包括概念验证研究、监管批准和现场应用。适当的评估方式可能因阶段而异，其衡量指标会从技术质量（质量控制）逐步扩展到社会影响 [1]。初期研究可能侧重于频率响应分析，而后续阶段则可能考虑与现有技术进行比较，该技术代表着临床相关应用中公认的金标准。与诊断效率相关的一种评估方法是采用基于任务的研究来评估图像质量，这种方法将技术质量指标扩展到特定的成像目的。

对目的的恰当描述应包括成像对象、图像的读取和分析方式，以及用于判定分析是否正确的衡量标准。对于诊断应用，图像可以由人类或计算机模型读取，后者包括统计模型和人工智能（AI）。

基于任务的评估由此能够实现对新技术的相关优化、对竞争技术的比较，以及对现有技术用于新应用的验证。若将其推至实用的极致，“技术”的定义可泛化为成像与

分析流程中的任何输入，其中既包括设备设计参数，也包括成像对象的特征、图像采集方案和阅片者资质。

人类阅片者与模型阅片者之间的比较，对于基于任务的评估本身的持续发展也至关重要。该领域的许多发展都源于医学成像应用，因此本文借鉴了相关研究中的术语。无论是旨在确立理想表现的上限，还是试图模仿人类，这些模型阅片者也都存在局限。与任务建模／选择及训练要求相关的问题可能会放大不确定性，最终降低评估的有效性。举例来说，理想观察者（IO）通常只能通过近似方式获得。此外，对于某些类别的可逆（或近似可逆）操作，理想观察者也不敏感，例如某些图像重建和滤波模式；这些操作对许多成像应用至关重要，也可能正是分析的对象。噪声的影响必须通过隐式建模来处理。

我们采用一种基于标量特征的数学仿真模型，这些特征受到外部和内部噪声源的共同影响。外部噪声源与不同案例之间的刺激（即待搜索的信号）变异性以及数据测量不确定性有关。内部噪声源则包括数据处理和非理想观察者行为所带来的不确定性。

我们的主要动机在于信号检测理论在医学成像试验中的应用。贝叶斯理想分类器是一种统计判别函数，通过受试者工作特征（ROC）曲线下面积来量化时，能够最大限度地提升许多诊断任务的性能。这一特性为优化受到成像物理、光子统计和患者特征所产生的噪声影响而性能下降的放射成像过程奠定了基础。另一项相关应用是通过虚拟成像试验评估和比较成像系统：在这类试验中，使用数学模型替代放射科医生，有望大幅加快有前景技术的临床应用[2]。其中一些模型通过为理想分类器（或更易构建的理想线性分类器）配备人类视觉的若干特性而构建，这些特性包括空间频率滤波器组，以及用于反映感知和认知不确定性的内部噪声。利用这些模型生成的ROC评

分数据本质上是图像特征的加权和，其评分方差同时反映了成像噪声和内部噪声。

我们探索能够执行整体视觉搜索（或概要）处理的新型统计观察者，部分方法是对可提取的最少图像特征设置阈值。在这些模型中，统计观察者减少了自由参数的数量，从而精简了系统。这些新型统计观察者可应用于虚拟成像试验中的放射科医生感知模型，以辅助其决策，并用于评估诊断成像工具。其他应用领域包括机器学习和计算机视觉，可用于性能基准测试以及特征选择与评估。这些模型还可用于国防/安全领域的目标检测与识别，以及传感器和探测器的评估。除了医学成像之外，工业和无损检测应用也需要稳健的统计观察者。

在第2节中，我们定义任务定义、ROC分析、理想观察者，以及截断如何影响分析。第3节介绍数学仿真结果，第4节则讨论截断观察者的优缺点。

2 理论

2.1 二元检测任务的ROC评估

二元任务将图像分为“阴性”（类别 $c = 0$）或“阳性”（$c = 1$）。在ROC研究中，观察者会为每幅图像提供一个标量评分统计量 λ，该统计量反映其对类别归属的信念（或置信度）。令 $p_\lambda(\lambda|c)$ 表示 λ的类别特异性概率分布。通过将评分与决策阈值 t_λ进行比较，可得出诊断结果。因此，这些评分会生成一条参数化ROC曲线，显示真阳性率和假阳性率（TPF和FPF）如何随阈值变化。任务表现由曲线下面积 [3]概括。

$$A_z = -\int_{-\infty}^{\infty} \mathrm{TPF}(t_\lambda)\, \frac{d\,\mathrm{FPF}(t_\lambda)}{dt_\lambda}\, dt_\lambda. \tag{1}$$

2.2 标准理想观察者分析

理想观察者根据图像特征计算 λ。可用的特征集 M 以向量 $\boldsymbol{f} = (f_1, \ldots, f_M) \in \mathbb{R}^M$表示。类别特定的联合特征分布记为 $p(\boldsymbol{f}|c)$，观察者通过根据对数似然比 [4]计算评分来最大化 A_z

$$\lambda(\boldsymbol{f}) = \ln \frac{p(\boldsymbol{f}|c=1)}{p(\boldsymbol{f}|c=0)}. \tag{2}$$

在这一理想情况下，特征变异性来自外部因素——类别归属和数据收集中的测量噪声。我们写作 $p(\boldsymbol{f}|c) = p_{\mathrm{ext}}(\boldsymbol{f}|c)$，以强调在不存在其他来源时这种噪声的来源。

观察者处理测得特征时的低效可视为会影响决策的内部噪声。这些低效是次优观察者的典型表现；我们的目标是研究处理观察者已识别出的低效的最优程序。本文考虑两种低效形式。其中一种情况是，对数似然

比是基于近似条件分布 $\hat{p}_{\text{ext}}(\boldsymbol{f}|c)$ 计算的。这些近似可能源于对 $p(\boldsymbol{f}|c)$ 进行参数估计时的不确定性，也可能是由于忽略了特定的分布成分。后一种情况的常见例子是在具有相关特征的任务中使用非预白化匹配滤波器。

第二种低效因素是测量后处理噪声，本文将其视为加性噪声，且认为它与特征和类别相互独立。在没有数据截断的情况下，我们的观察者会使用多元概率分布

$$p(\boldsymbol{f}|c) = p_{\text{ext}}(\boldsymbol{f}|c) * p_{\text{int}}(\boldsymbol{f}) \tag{3}$$

这些分布由测量分布和处理分布经过 M 维卷积得到。根据这些假设，内部噪声分布 p_{int} 被认为是可分离的。

2.3 截断特征分布

数据截断使理想观察者明确进行特征提取，将两种噪声过程分开，使内部噪声能够惩罚对非信息性数据的使用。本文考虑使用单位高度阶跃函数 step($f_i - \tau_i$) 实现的左侧（或下侧）截断；该函数保留 $f_i \geq \tau_i$ 的特征值，适用于 $i = 1$、.....、M。右侧截断和双侧截断则是调整决策中对真阳性和真阴性相对重视程度的其他选项。左侧阈值集合归入向量 $\boldsymbol{\tau} = (\tau_1, ..., \tau_M)$。一种典型的计算目标是根据 ROC面积优化该向量；考虑到涉及的组合维度，人类观察者能否完成这些计算尚不明确。本文中的示例模拟仅对一维和二维阈值进行扫描。

这些阈值将特征空间划分为 2^M 个互不相交的区域，每个区域都为观察者定义了一个独特的提取特征子集。我们将每个子集与一个含有 α_i 的 M维指示向量 $\boldsymbol{\alpha}$ 相关联，其中提取的特征和

被拒绝的特征索引分别对应值 1 和 0。空集对应于所有特征均被拒绝且给定图像未获评分的事件，其 $\alpha_i = 0$ 对于所有 i均成立，记作 $\boldsymbol{\alpha} = \mathbf{0}$。完整提取的情形为 $\boldsymbol{\alpha} = \mathbf{1}$。在左侧截断下，与给定 $\boldsymbol{\alpha}$ 相关联的区域具有如下特征支撑集 $\boldsymbol{\chi}_\alpha(\boldsymbol{\tau}) = \{\boldsymbol{f}\colon\ S_\alpha(\boldsymbol{f}, \boldsymbol{\tau}) = 1\}$，如复合阶跃函数所规定

$$S_\alpha(\boldsymbol{f}, \boldsymbol{\tau}) = \prod_{i=1}^{M} \left[1 - \mathrm{step}(f_i - \tau_i)\right]^{1-\alpha_i} \left[\mathrm{step}(f_i - \tau_i)\right]^{\alpha_i}. \tag{4}$$

补集 $\{\boldsymbol{f}\colon\ S_\alpha(\boldsymbol{f}, \boldsymbol{\tau}) = 0\}$ 定义了 $\boldsymbol{\alpha}$ 在右侧截断下的支持集。

采用这种支持集表示法，$\boldsymbol{\alpha}$-提取概率——即提取某一特定特征子集的概率——可表示为

$$Pr_{\mathrm{ext},\tau}(\boldsymbol{\alpha}|c) = \int_{\boldsymbol{\chi}_\alpha(\boldsymbol{\tau})} p_{\mathrm{ext}}(\boldsymbol{f}|c)\, d\boldsymbol{f}. \tag{5}$$

对于独立特征，这些概率是各个特征的拒绝概率与互补接受概率的乘积。对于一幅图像，拒绝所有特征的概率也可计算为

$$Pr_{\mathrm{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c) = 1 - \sum_{\boldsymbol{\alpha} \neq \mathbf{0}} Pr_{\mathrm{ext},\tau}(\boldsymbol{\alpha}|c), \tag{6}$$

这将在我们的ROC分析中得到特别考量。

描述非零 $\boldsymbol{\alpha}$ 的特征分布，需要将 $\boldsymbol{f}$ 中的元素分成一对变长向量 $(\boldsymbol{e}, \boldsymbol{e}')$，其中 $\boldsymbol{e}$ 包含 m 提取出的元素（f_i 满足 $\alpha_i = 1$），而其补集 $\boldsymbol{e}'$ 包含被 M - m 拒绝的元素（f_i 满足 $\alpha_i = 0$）。在具有完整特征信息的情况下（即 $m = M$），其分布为

$$p_{\mathrm{ext},\tau}(\boldsymbol{e}|\boldsymbol{\alpha}, c) = \frac{1}{Pr_{\mathrm{ext},\tau}(\boldsymbol{\alpha}|c)}\, p_{\mathrm{ext}}(\boldsymbol{f}|c)\ S_\alpha(\boldsymbol{f}, \boldsymbol{\tau}). \tag{7}$$

否则，该分布通过边缘计算得到

$$p_{\text{ext},\tau}(\boldsymbol{e}|\boldsymbol{\alpha},c) = \frac{1}{P_{\text{ext},\tau}(\boldsymbol{\alpha}|c)} \int_{\mathbb{R}^{M-m}} p_{\text{ext}}(\boldsymbol{f}|c)\ S_\alpha(\boldsymbol{f},\boldsymbol{\tau})\, d\boldsymbol{e}'. \tag{8}$$

我们将 $p_{\text{ext},\ \tau}(\boldsymbol{e}|\boldsymbol{\alpha},c)$称为截断外部源分布。请注意，无需为 $\boldsymbol{\alpha}=\mathbf{0}$ 设置分布分量，因为全部特征拒绝的影响将在ROC面积公式中通过单独一项处理。

数据模型的最后一步是按照公式3的方式通过卷积加入内部噪声，得到

$$p_\tau(\boldsymbol{e}|\boldsymbol{\alpha},c) = p_{\text{ext},\tau}(\boldsymbol{e}|\boldsymbol{\alpha},c) * p_{\text{int}}(\boldsymbol{e}), \tag{9}$$

但需注意，$p_{\text{ext},\ \tau}(\boldsymbol{e}|\boldsymbol{\alpha},c)$是 m 特征的函数，且卷积仅应用于该特征子集。

2.4 截断数据的ROC分析

优化阈值时的一般权衡显而易见：既要控制内部噪声的起始，又要丢弃可能有用的数据。多变量数据在一定程度上减轻了截断的负面影响，因为只保留一个特征值也能获得评分。然而，即使是适度的截断也会导致大量ROC数据损失；三个接受概率均为$_{1\ 2}$ 且相互独立的特征会导致八分之一的图像成为未评分案例。在ROC分析中如何处理未评分案例，会显著影响优化 [5]。

–

不完整的分析会在真阳性和假阳性计数中排除未评分案例，但在计算 TPF 和 FPF 时仍将其保留在分母中。采用左侧截断时，这会产生一条在以下坐标处终止的部分 ROC曲线

$$(\text{TPF},\text{FPF}) = (1 - Pr_{\text{ext},\tau}(\boldsymbol{\alpha}=\mathbf{0}|c=1), 1 - Pr_{\text{ext},\tau}(\boldsymbol{\alpha}=\mathbf{0}|c=0)). \tag{10}$$

如第2.3节所述，$Pr_{\text{ext},\ \tau}$ 这些项是两个图像类别中完全拒绝特征的概率。为了优化 $\boldsymbol{\tau}$，部分曲线对 A_z造成的显著负面影响（见图1）往往会盖过内部噪声的影响；在这种情况下，最优方案是通过减少截断来补全 ROC曲线。

有意义的优化可以补全曲线。在左侧截断的情况下，如何做到这一点，在强迫选择研究的背景下显而易见。观察者的任务是从一对图像中识别出阳性案例，而 A_z 是成功的概率。图 1 中阴影区域的面积$A_z^{(1)}$表示图像对被评分且被正确标记的联合概率。我们将其视为观察者表现的分析成分。图 1 中由水平虚线延伸段界定、面积为

$$A_z^{(2)} = [1 - Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 1)]\ Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 0), \tag{11}$$

源于将未评分的阴性案例与已评分的阳性图像进行比较。从 ROC 的角度来看，这是通过如下修改假阳性分数（FPF）得出的：

$$\Delta FPF = Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 0)\,\delta(t_\lambda + \infty), \tag{12}$$

通过预先指定（借助狄拉克 δ函数）未评分的阴性图像的最低评分 $\lambda = -\infty$ 来实现。将未评分案例与事实上被归为阴性分类联系起来后，我们将 $A_z^{(2)}$称为观察者表现的概要成分。

未评分的阳性图像与已评分的阴性图像进行强迫选择比较，并不会增加 A_z。不过，观察者对两张未评分图像进行猜测也会带来第三项表现贡献；在二元任务中，猜中的几率相等。这一猜测成分通过一段点状直线段将ROC曲线延伸至点(1,1)，如图1所示。三角形面积

$$A_z^{(3)} = \frac{1}{2}\,Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 1)\,Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 0) \tag{13}$$

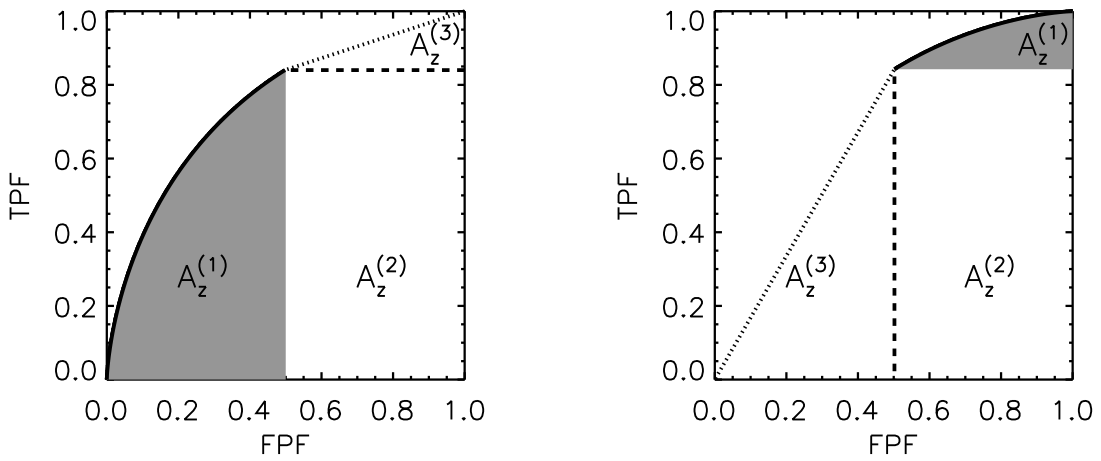

图1：使用所提出的截断分析处理未评分的ROC案例。（a）在左侧截断的情况下，部分曲线（以实线表示）产生了由阴影区域表示的缩减面积$A_z^{(1)}$。为所有未评分的阴性案例分配最低评分（$\lambda = -\infty$）可通过使用水平虚线补全曲线，提供额外面积 $A_z^{(2)}$。这种补全不允许为猜测提供增益，而为未评分正例同样分配最低评分，则会产生以点线表示的猜测扩展，其面积增量为 $A_z^{(3)}$。（b）基于右侧截断的ROC曲线补全。ROC曲线使用单一特征计算。

也与公式12中的假阳性分数增量以及真阳性分数增量相关

$$\Delta TPF = Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 1)\,\delta(t_\lambda + \infty). \tag{14}$$

因此，概要和猜测成分 $A_z^{(2)}$和 $A_z^{(3)}$由外部来源的概率分布和 $\boldsymbol{\tau}$决定。底层的观察者模型或内部噪声不会产生影响，尽管它们是决定截断如何影响表现的分析成分的因素。在其他条件相同的情况下，纳入猜测成分的优化，与忽略猜测成分的优化相比，往往会采用更大程度的数据截断。若数据截断使 A_z增加，则可将这种优化解读为：以降低处理更模糊案例时的表现为代价，提升处理相对简单案例时的表现（与不截断相比）。

2.5 使用截断数据的理想观察者

我们的理想观察者使用适用于给定图像数据的对数似然比来计算评分，将公式 2 推广为

$$\lambda_{\alpha,\tau}(\boldsymbol{e}) = \ln \frac{p_\tau(\boldsymbol{e}|\boldsymbol{\alpha}, c=1)}{p_\tau(\boldsymbol{e}|\boldsymbol{\alpha}, c=0)}. \tag{15}$$

由公式 15 推导出的类条件评分分布表示为 $p_{\lambda,\alpha,\tau}(\lambda|c)$。总体而言，该模型从混合分布中生成评分

$$p_{\lambda,\tau}(\lambda|c) = \sum_{\boldsymbol{\alpha}\neq\mathbf{0}} Pr_{\text{ext},\tau}(\boldsymbol{\alpha}|c)\, p_{\lambda,\alpha,\tau}(\lambda|c). \tag{16}$$

仅根据已评分图像计算得到的相关 ROC 量为

$$\text{TPF}(t_\lambda) = \sum_{\boldsymbol{\alpha}\neq\mathbf{0}} Pr_{\text{ext},\tau}(\boldsymbol{\alpha}|c=1) \int_{t_\lambda}^{\infty} p_{\lambda,\alpha,\tau}(\lambda|c=1)\, d\lambda \tag{17}$$

$$\frac{d\,\text{FPF}(t_\lambda)}{dt_\lambda} = \sum_{\boldsymbol{\alpha}\neq\mathbf{0}} Pr_{\text{ext},\tau}(\boldsymbol{\alpha}|c=0)\, p_{\lambda,\alpha,\tau}(t_\lambda|c=0). \tag{18}$$

根据公式 1，并纳入与未评分图像相关的面积分量 $A_z^{(2)}$和 $A_z^{(3)}$，ROC性能为

$$A_z = A_z^{(1)} + \left[1 - \frac{1}{2} Pr_{\text{ext},\tau}(\boldsymbol{\alpha}=\mathbf{0}|c=1)\right] Pr_{\text{ext},\tau}(\boldsymbol{\alpha}=\mathbf{0}|c=0) \tag{19}$$

与评分推导出的分量

$$A_z^{(1)} = \sum_{\boldsymbol{\alpha},\tilde{\boldsymbol{\alpha}}\neq\mathbf{0}} Pr_{\text{ext},\tau}(\boldsymbol{\alpha}|c=1)\, Pr_{\text{ext},\tau}(\tilde{\boldsymbol{\alpha}}|c=0)\, A_{z,\alpha,\tilde{\alpha},\tau}^{(1)}, \tag{20}$$

作为交叉-$\boldsymbol{\alpha}$ 面积贡献的加权和

$$A^{(1)}_{z,\alpha,\tilde{\alpha},\tau} = -\int_{-\infty}^{\infty} p_{\lambda,\tilde{\alpha},\tau}(t_\lambda|c=0) \int_{t_\lambda}^{\infty} p_{\lambda,\alpha,\tau}(\lambda|c=1)\, d\lambda\, dt_\lambda. \tag{21}$$

使用未经截断且不含内部噪声的评分时，理想观察者会生成一条非凹ROC曲线，为包含线性线段的曲线提供上界（见图2）。在没有截断且内部噪声幅度固定的情况下，该观察者为使用相同内部噪声幅度并经截断得到的ROC曲线提供下界。截断条件下的最优曲线 $\boldsymbol{\tau}$ 与上界曲线相交于由式10给出的阈值工作点，这表明，如果能够利用经过最少处理的提取特征进行决策，从而避免后续的内部噪声来源，就能实现理想表现。交点是 $\boldsymbol{\tau}$的函数；在第3节的模拟中，通过最大化 A_z来确定最优的 $\boldsymbol{\tau}$ 。

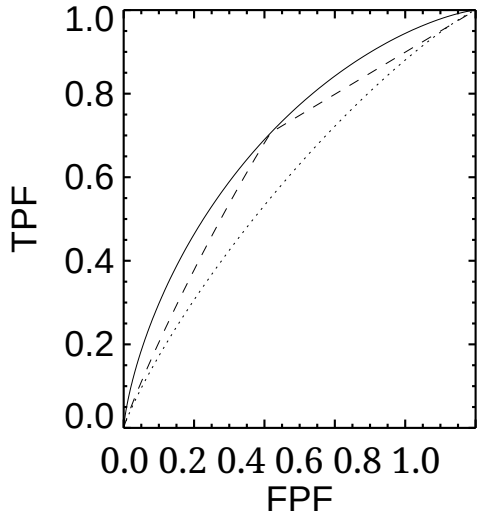

图 2：理想观察者和截断理想观察者的 ROC曲线。考虑了单一特征。点线和实线分别表示存在和不存在方差为 σ^2 的内部噪声时的理想观察者。虚线表示具有相同内部噪声方差的阈值化模型。

2.6 高内部噪声下的理想观察者

对于任何观察者模型，在高内部噪声极限下，分量 $A^{(1)}_{z\ \alpha}{}_{,,\tilde{\alpha},\tau}$ 的值为 $_{1\ 2}$，而对于采用截断_评分的理想观察者，曲线下面积作为 $\boldsymbol{\tau}$ 的函数为

$$A_z = \frac{1}{2}\left[1 + Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 0) - Pr_{\text{ext},\tau}(\boldsymbol{\alpha} = \mathbf{0}|c = 1)\right]. \tag{22}$$

下一节将基于这一渐近结果以及具有 $M \in \{1,2\}$个特征的截断理想观察者的其他性质，考察优化问题。

3 计算机模拟

我们考虑正态分布特征，并采用记号 $p_{\text{ext}}(\boldsymbol{f}|c) = \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)$ 表示外部效应，其中 $\boldsymbol{\mu}_c$ 和 $\boldsymbol{\Sigma}_c$ 分别表示类别 c的特征均值向量和 $M \times M$协方差矩阵。特征 f_i 的类别均值为 $\mu_{c,i}$ ，类别方差为$\sigma^2_{c,i}$。我们测试独立特征模型和相关特征模型。所有试验均采用正态分布的零均值内部噪声，因此 $p_{\text{int}}(\boldsymbol{f}) = \mathcal{N}(0, \sigma^2\boldsymbol{I})$，其方差为 σ^2，而 $\boldsymbol{I}$为 M维单位矩阵。这一特定的阈值观察者模型包含 M 个特征阈值、 $2M$ 个特征均值、 $2M + 1$ 个特征方差以及 $M^2 - M$个特征协方差参数。对于独立特征，不存在协方差。

3.1 单特征试验

概率分布塑造的一些基本性质可通过 $M = 1$ 清楚地看出。我们将标准外部源分布 $\mathcal{N}(0, 1)$ 分配给负类，并测试正类变体 $\mathcal{N}(0.75, \sigma_1^2)$，使得 σ_1^2 小于、等于或大于 1。在没有截断的情况下，理想观察者在同方差情形下是线性函数，否则为二次函数。引入截断后，同方差模型变为分段线性模型，同时向二次模型中加入高阶项。理想观察者的完全信息包含在公式 9 的处理后的分布中，

对于单个或独立正态特征，其闭式表达式为

$$p_\tau(\boldsymbol{e}|\boldsymbol{\alpha}, c) = \prod_{\substack{i=1 \\ \alpha_i=1}}^{M} \mathcal{N}(\mu_{c,i}, \sigma_{c,i}^2 + \sigma^2)\Phi(\frac{f_i - a_{c,i}}{b_{c,i}}) \,/\Phi(\frac{\mu_{c,i} - \tau_i}{\sigma_{c,i}}), \tag{23}$$

当 Φ(·)时 ，标准正态累积分布函数（CDF）和参数 s

$$a_{c,i} = \frac{1}{\sigma_{c,i}^2} \left[\tau_i \left(\sigma_{c,i}^2 + \sigma^2\right) - \sigma^2 \mu_{c,i}\right] \tag{24}$$

$$b_{c,i} = \frac{\sigma}{\sigma_{c,i}} \sqrt{\sigma_{c,i}^2 + \sigma^2}. \tag{25}$$

随着内部噪声降低（即 $\sigma \to 0$），CDF 趋近于在 $f_i = \tau_i$处存在不连续点的阶跃函数极限。

图 3 中的上排图表证实了 $\sigma = 0$时未截断性能的最优性。如红色曲线所示，未正则化性能随 τ 单调下降。正则化模型（黑色曲线）下扩展稳定性的程度，取决于概要处理过程（蓝色曲线）仅截断负例的有效性，从而防止与正例截断增加相关的猜测（绿色曲线）。对于 $\sigma_1 = 3$，较低 τ 下的不稳定性源于对正例的优先截断（猜测成分未能充分体现这一点）。随后出现的结果是：当 τ 的取值介于类别均值之间时，表现出现次级峰值。

如图 3 下排图表所示，随着内部噪声增加，这种阈值峰值对所有 σ_1都显现出其重要性。由于概要处理更有效，较小的 σ_1 似乎可能带来更大的阈值优势，但至少依据一种相关指标，这种效应在各个 σ_1中是一致的。我们将最小截断效率定义为 σ 的截断性能与未截断理想观察者表现之比。这种效应在内部噪声较高时最为显著；此时 PDS 可能带来的最大

收益体现在：极端内部噪声下的最优观察者性能，显著高于理想观察者加入内部噪声后得到的 0.5。即使没有信息性评分（即方差为无穷大），最优性能仍高于偶然水平（50%）。

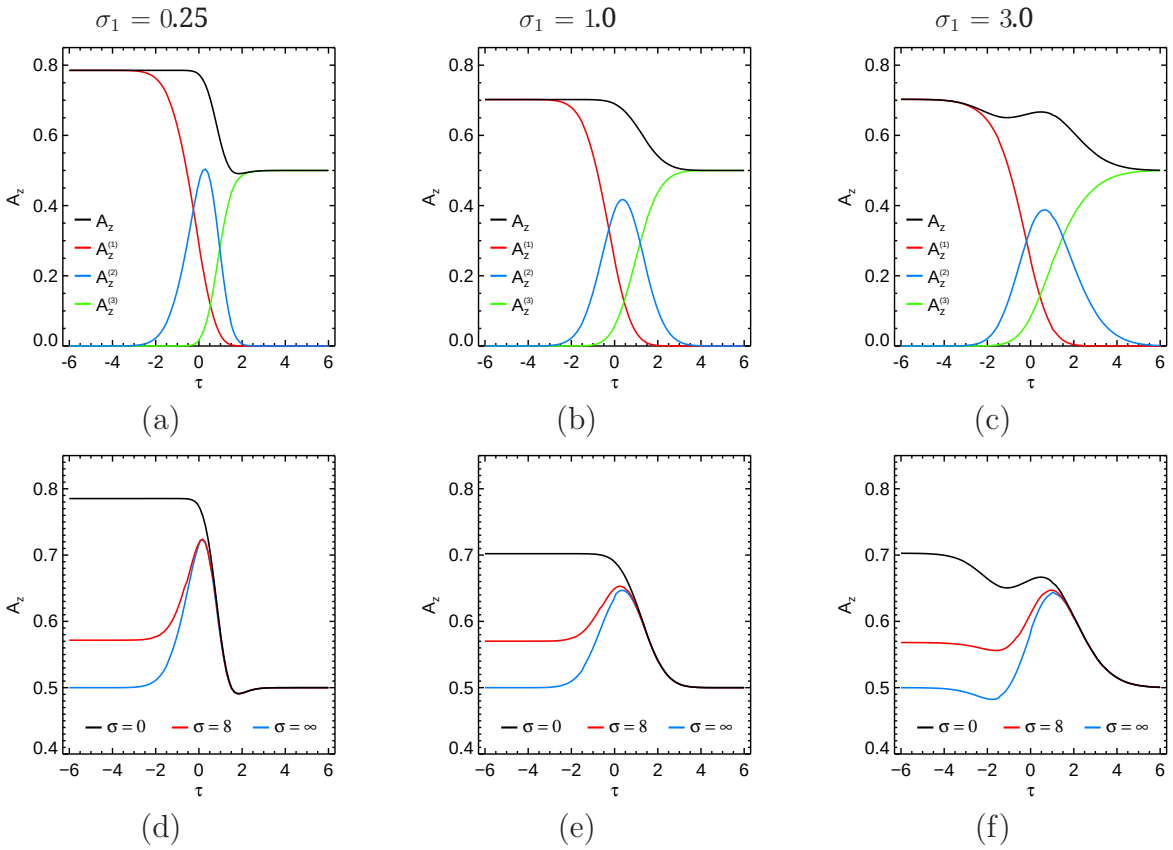

图3：三种单特征模拟的截断模型表现比较。在这些模拟中，正态特征遵循标准分布，而异常特征的均值为0.75。从左到右，异常特征的标准差分别为0.25、1.0和3.0。上排（图a-c）显示了无内部噪声时的整体表现（黑色），并展示了分析（红色）、概要（蓝色）和猜测（绿色）的贡献。下排（图d-f）显示了整体表现随内部噪声阈值 $\sigma \in \{0, 8, \infty\}$变化的情况。

3.2 双特征试验

当 M 大于1时，特征之间的相互作用会影响观察者表现。S

通过具有两个独立特征的研究来说明这一点。

我们为负类指定标准外部源分布 $\mathcal{N}(0,1)$，并测试正类变体 $\mathcal{N}(0.75,\sigma_1^2)$，使得 σ_1^2 小于、等于或大于 1。在未进行截断时，同方差情形下的理想观察者是线性函数，其他情形下则是二次函数。引入截断后，同方差模型变为分段线性模型，而二次函数模型则增加了高阶项。

理想观察者的完全信息包含在式 9 的处理后分布中；对于单个或相互独立的正态特征，其闭式表达式为

$$p_\tau(\boldsymbol{e}|\boldsymbol{\alpha},c)=\prod_{\substack{i=1\\ \alpha_i=1}}^{M}\mathcal{N}(\mu_{c,i},\sigma_{c,i}^2+\sigma^2)\Phi(\frac{f_i-a_{c,i}}{b_{c,i}})\,/\Phi(\frac{\mu_{c,i}-\tau_i}{\sigma_{c,i}}), \tag{26}$$

其中 $\Phi(\cdot)$ 是标准正态累积分布函数（CDF），参数为

$$a_{c,i}=\frac{1}{\sigma_{c,i}^2}\left[\tau_i\left(\sigma_{c,i}^2+\sigma^2\right)-\sigma^2\mu_{c,i}\right] \tag{27}$$

$$b_{c,i}=\frac{\sigma}{\sigma_{c,i}}\sqrt{\sigma_{c,i}^2+\sigma^2}. \tag{28}$$

在额外假设各特征同分布的情况下，无论 M 如何，都只需要一个阈值。图 4 展示了阈值、处理噪声方差和特征数量对理想观察者性能的影响。图中显示，当截断数据的 $\boldsymbol{\tau}$ 接近双类别均值时，会出现局部最大值。

理想观察者对特征的截断敏感，而对特征值的变化不敏感。通过采用调整得当的阈值，内部噪声水平不同的观察者也能取得相似的表现；不过，内部噪声会加剧阈值调整不当所带来的损失。

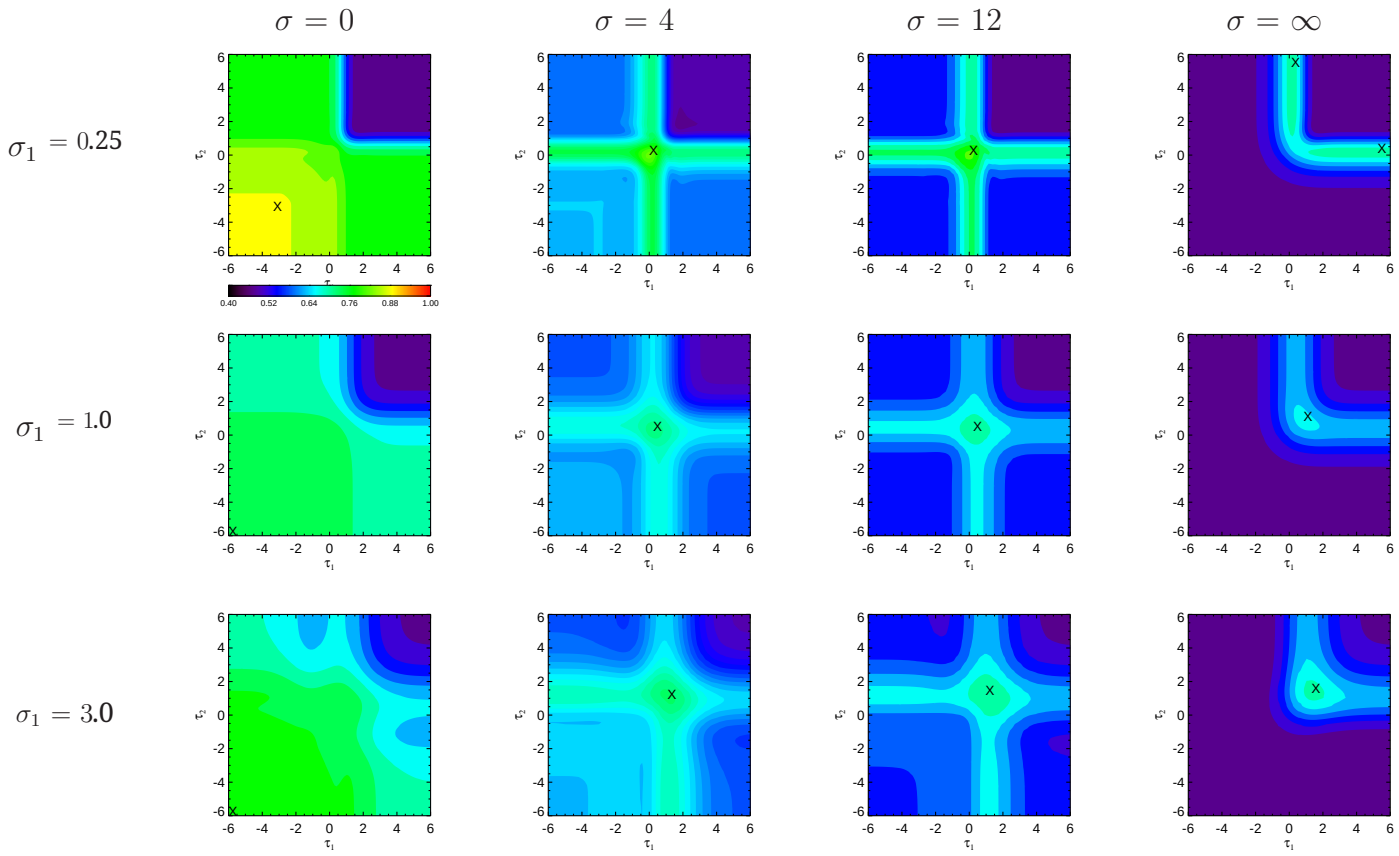

图4：双特征试验的结果。x 标记表示数据截断时，在给定噪声方差（由 sigma 表示）下理想观察者表现的最大值。x 轴和 y 轴分别表示各变量的取值范围，其中 0 表示分布的中点。外部噪声由 $\sigma 1$ 表示，其正态分布由 $\sigma 1{=}1$ 给出

4 讨论

我们为噪声排除模型引入了一种理想观察者。给定区域的数据会被固定的下阈值截断。对于截断数据，二元理想观察者的 AUC 由分析贡献、概要贡献和猜测贡献构成。在典型二选一模型中，分析贡献来自对有评分的正常案例与异常案例的比较；概要成分来自对一个有评分案例和一个无评分案例的比较；猜测成分则来自对无评分的正常案例与异常案例的比较。

当 a 较低时 在噪声量方面，分析成分是主要贡献因素，而 AUC

在截断程度较低时达到最高。随着噪声增大，概要处理过程成为主要成分。噪声过大时，猜测也很重要。

将 IO 视为相较于不含内部噪声的传统理想观察者而言的次优观察者时，我们有必要了解传统理想观察者需要多大的加性内部噪声，才能达到次优观察者的表现。

5 致谢

本研究由美国国家生物医学成像与生物工程研究所资助，资助编号为 R01EB032416。本文内容完全由作者负责，不一定代表美国国家生物医学成像与生物工程研究所的官方观点。

参考文献

[1] D. G. Fryback 和 J. R. Thornbury，“诊断成像的疗效。”《医疗决策》，第 11 卷，第 88–94 页，1991 年 4 月。

[2] A. Badano、C. G. Graff、A. Badal、D. Sharma、R. Zeng、F.W. Samuelson、S. J. Glick 和 K. J. Myers，“通过计算机模拟成像试验评估数字乳腺断层合成替代全视野数字乳腺摄影的可行性。”《JAMA网络开放》，第 1 卷，e185474 页，2018 年 11 月。

[3] H. H. Barrett、C. K. Abbey 和 E. Clarkson，《图像质量的客观评估。III. ROC 指标、理想观察者和似然生成函数》，J. Opt. Soc. Am. A，第 15 卷，第 1520–1535 页，1998 年。

[4] W. S. Geisler，《理想观察者分析》。2003 年。

[5] Z. Karbaschi 和 H. C. Gifford，《利用视觉搜索模型观察者评估 CT采集参数》，《医学成像杂志》，第 5 卷，第 2 期，页码 025501，2018 年。