

# 消化粘膜下超声图像斑点的模型分析

伍一渐

上海理工大学理学院, 上海

收稿日期: 2023年12月5日; 录用日期: 2023年12月25日; 发布日期: 2024年2月29日

## 摘要

由于超声斑点的形成依赖于底层组织结构, 利用斑点建模对超声图像中的组织器官进行分割和分类已经成为一个研究热点。本文将探究最能表征消化粘膜下正常组织和肿瘤组织中斑点的概率分布模型。本文对消化粘膜下正常组织和肿瘤组织斑点的一阶灰度值进行建模。模型是通过分析从上海市第六人民医院消化内镜中心提供的197张消化道粘膜下肿瘤超声图像取而得到的。在消化粘膜下超声图像中手动截取了正常组织和肿瘤组织图像, 将其灰度直方图作为斑点分布的概率密度函数, 使用极大似然估计来拟合斑点分布, 并利用K-S检验和 $R^2$ 检验来评估超声斑点与五种概率分布之间的拟合程度。在分析消化黏膜正常组织的直方图时, 伽玛分布显示出比其他概率分布模型更好的适用性, 具有最小的K-S统计值为0.0065和最高的R方值为0.9967。而在分析消化黏膜肿瘤组织的直方图时, 威布尔分布相对于其他模型更合适, 显示出最小的K-S统计值为0.0096和最大的R方值为0.9956。因此伽马分布更适合描述消化道粘膜下正常组织的斑点特征, 而威布尔分布更适合描述消化道粘膜下肿瘤组织的斑点特征。

## 关键词

消化道粘膜下肿瘤, 超声图像, 斑点分析, 伽马分布, 威布尔分布

# Model Analysis of Ultrasound Image Speckles in the Digestive Submucosa

Yijian Wu

College of Science, University of Shanghai for Science and Technology, Shanghai

Received: Dec. 5<sup>th</sup>, 2023; accepted: Dec. 25<sup>th</sup>, 2023; published: Feb. 29<sup>th</sup>, 2024

## Abstract

Since the formation of ultrasound speckles depends on the underlying tissue structure, segmentation and classification of tissues and organs in ultrasound images using speckle modeling has be-

come a research hotspot. In this paper, we will explore the probability distribution model that best characterizes speckle in digestive submucosal normal and tumor tissues. In this paper, the first-order gray values of normal and tumor tissue speckles under the digestive mucosa are modeled. The model was obtained by analyzing 197 ultrasound images of digestive submucosal tumors taken from the Gastrointestinal Endoscopy Center of Shanghai Sixth People's Hospital. Normal and tumor tissue images were manually intercepted in the digestive submucosal ultrasound images, and their grayscale histograms were used as the probability density function of the speckle distributions, and great likelihood estimation were used to fit the speckle distributions, and the K-S test and the test were used to assess the degree of fit between the ultrasound speckles and the five probability distributions. In analyzing histograms of digestive mucosal normal tissue, the gamma distribution showed better fit than the other probability distribution models, having the smallest K-S statistic value of 0.0065 and the highest R-squared value of 0.9967. whereas, in analyzing histograms of digestive mucosal tumor tissue, the Weibull distribution was a better fit relative to the other models, showing the smallest K-S statistic value of 0.0096 and the largest R-squared value of 0.9956. Thus Gamma distribution is more suitable to describe the speckle characteristics of normal tissues under the digestive mucosa whereas Weibull distribution is more suitable to describe the speckle characteristics of tumor tissues under the digestive mucosa.

## Keywords

Digestive Tract Submucosal Tumors, Ultrasound Images, Speckle Analysis, Gamma Distribution, Weibull Distribution

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近年来,随着大众健康意识的增强、内镜检查的普及以及超声内镜技术的不断发展与完善,消化道粘膜下肿瘤(SMT)的检出率逐年上升。超声内镜[1](EUS)具备超声和内镜双重检查的优势,可以通过内镜清晰地观察粘膜下肿瘤的外部形态,并通过小探头超声对肿瘤进行扫描,以此获得超声图像。当超声波脉冲在组织中传播并与组织相互作用时,脉冲在多个方向上散射,发生干涉现象,形成散斑。由于生物软组织的密度、弹性等的不均匀性,超声波在人体内不同组织的分布可能存在差异[2]。由于散斑的形成依赖于组织,因此对组织的散斑进行建模可以更好地得到组织结构或与区域形态的相关信息,对超声斑点的分析还可以得到额外的非侵入性信息,以此了解体内组织中疾病的存在和类型[3]。Zhang 等人[4]提出医学超声图像中散斑的分布更加倾向于广义的伽马分布。边海丽[5]从医学超声图像中散斑产生的原理出发,推理得出医学超声图像中散斑分布为瑞利分布。Tsui 等人[6]应用斑点模型与活检相比较来表征肝脏疾病。同样是肝脏,Wan 等人[7]调查了 107 例患者,发现器官脂肪含量与 Nakagami 分布参数之间存在显著的正相关。Nasr 等人[8]使用超声图像区分刚切除的和脱细胞的啮齿动物的肾脏。这些工作表明,量化组织结构的变化,并将它们与斑点特征联系起来是有可能的。消化粘膜下组织的细微变化,无论是区分组织还是检测疾病,都可能反映在斑点上。这样的特征助于肿瘤和正常组织的区分。然而,哪种统计分布能最好地模拟散斑还有待观察。

在本文中,我们将理论和实践相结合,从实际的消化粘膜下组织超声图像出发,使用 143 个临床患者的真实超声数据,将不同组织斑点的灰度直方图作为概率密度函数,评估五种概率密度模型,得出更符合实际的消化粘膜下肿瘤和正常消化粘膜的超声图像斑点分布的模型。

## 2. 研究方法

### 2.1. 极大似然估计

极大似然估计法是一种常用的统计推断方法，通常被用于根据数据构建概率模型。其基本思想是通过优化对概率分布的似然函数，找到最大化该函数的参数值，从而得到最适合数据的概率模型。它的好处在于提供了一种简单而高效的方法，从给定的数据集中估计出一个恰当的概率模型。它不仅可以处理离散和连续的数据，而且对于大多数常见的分布，都有相应的最优参数值。

假设有一个样本集合  $\{x_1, x_2, \dots, x_n\}$ ，它们服从某种概率分布。我们要估计这种分布的参数，使得给定数据集合的条件下，概率分布的似然函数最大化。极大似然估计如表 1 所示。

Table 1. Maximum-likelihood estimate

表 1. 极大似然估计

实验步骤
1) 找到似然函数
$L(\theta) = p(x_1, x_2, \dots, x_n   \theta) = \prod_{i=1}^n p(x_i   \theta)$
2) 对似然函数取对数
$H(\theta) = \ln(L(\theta)) = \ln(p(x_1, x_2, \dots, x_n   \theta)) = \sum_{i=1}^n \ln(p(x_i   \theta))$
3) 对 $H(\theta)$ 求极大值:
$\frac{d}{d\theta} H(\theta) = \frac{d}{d\theta} \ln(L(\theta)) = 0$
4) 得到参数
$\hat{\theta} = \arg \max_{\theta} H(\theta)$

### 2.2. K-S 检验

Table 2. Kolmogorov-Smirnov test

表 2. Kolmogorov-Smirnov 检验

实验步骤
1) 计算样本的经验分布函数(ECDF)的值 $F_n(X)$ ，其中 $N$ 是样本大小， $n \in N$
$F_n(X) = \frac{1}{n} \sum_{i=1}^n I_{-\infty}^n(x_i)$
2) 计算理论分布的 CDF 值 $G_n(X)$ 。
3) 计算 $F_n(X)$ 和 $G_n(X)$ 两个累积分布函数的最大差距 $D_n$
$D_n = \max_{n \in N}  F_n(X) - G_n(X) $
4) 确定显著性水平(通常为 0.05)，表示对原假设的拒绝程度。
5) 判断原假设
如果 K-S 统计量小于临界值，则接受原假设，认为样本与理论分布之间没有显著差异；
如果 K-S 统计量大于临界值，则拒绝原假设，认为样本与理论分布存在显著差异。

K-S (Kolmogorov-Smirnov) 检验是一种用于检验样本分布与理论分布是否一致的统计方法。该方法广泛应用于各种分布检验和统计分析中，可以用于检验各种常见分布的拟合情况。该方法通过计算两个累

积分布函数(CDF)之间的最大距离来评估两个分布之间的差异。

假设有一个来自总体分布  $F(X)$  的样本集合  $\{x_1, x_2, \dots, x_n\}$ 。要检验样本是否来自由  $G(X)$  定义的某一理论分布。K-S 检验[9]如表 2 所示。

### 2.3. 拟合优度检验

拟合优度(Goodness of Fit)是指回归直线对观测值的拟合程度。度量拟合优度的统计量是可决系数(亦称确定系数)  $R^2$ 。它表示由模型解释的数据方差占总数据方差的比例,通常取值范围在 0 到 1 之间。 $R^2$  的表达式为:

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (1)$$

其中 RSS 为残差平方和, TSS 为总平方和。而  $R^2$  取值越接近 1, 说明模型的拟合效果越好。 $R^2$  取值越接近 0, 说明模型的拟合效果越差。

## 3. 模型

本文选择了五种参数分布模型作为消化粘膜下肿瘤和正常消化粘膜灰度水平的潜在模型。

### 3.1. 伽马分布

伽马分布是由两个参数控制的连续概率分布,两个参数分别为形状参数  $\alpha > 0$  和尺度参数  $\beta > 0$ 。伽马分布的概率密度函数模型[10]表示为:

$$p_\gamma(x|\alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2)$$

### 3.2. 对数正态分布

对数正态分布是由两个参数控制的连续概率分布,两个参数分别为对数均值  $\mu \in (-\infty, +\infty)$  和对数标准差  $\sigma > 0$ 。对数正态分布的概率密度函数模型[11]表示为:

$$p_{\ln}(x|\mu, \sigma) = \begin{cases} \frac{1}{x\sqrt{2\pi}\sigma} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (3)$$

### 3.3. 正态分布

正态分布是由两个参数控制的连续概率分布。两个参数分别为均值  $\mu \in (-\infty, +\infty)$  和标准差  $\sigma > 0$ 。它的概率密度函数模型[12]表示为:

$$p_n(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

### 3.4. 瑞利分布

瑞利分布是由一个参数控制的连续概率分布。这个参数为尺度  $\sigma > 0$ 。参数它的概率密度函数模型[13]表示为:

$$P_{rayl}(x|\sigma) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (5)$$

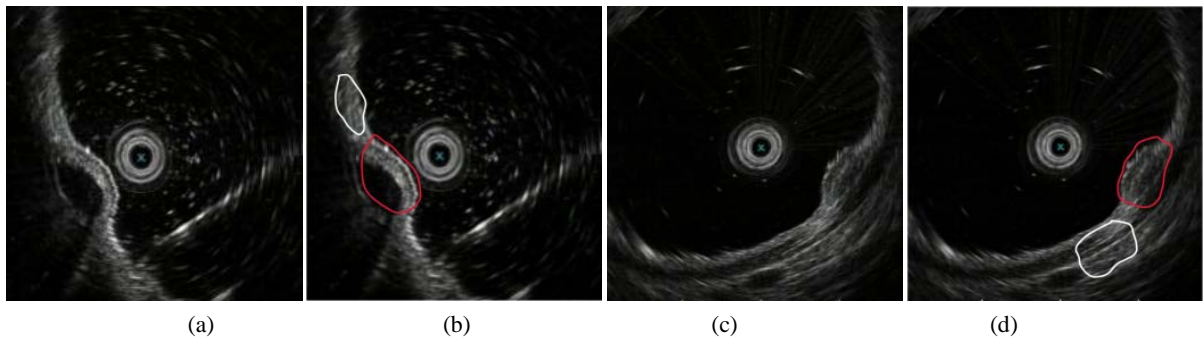
### 3.5. 威布尔分布

布威布尔分布是由两个参数控制的连续概率分布，两个参数分别为形状参数  $\alpha > 0$  和尺度参数  $\beta > 0$ 。它的概率密度函数模型[14]表示为：

$$P_{wbl}(x|\alpha, \beta) = \begin{cases} \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

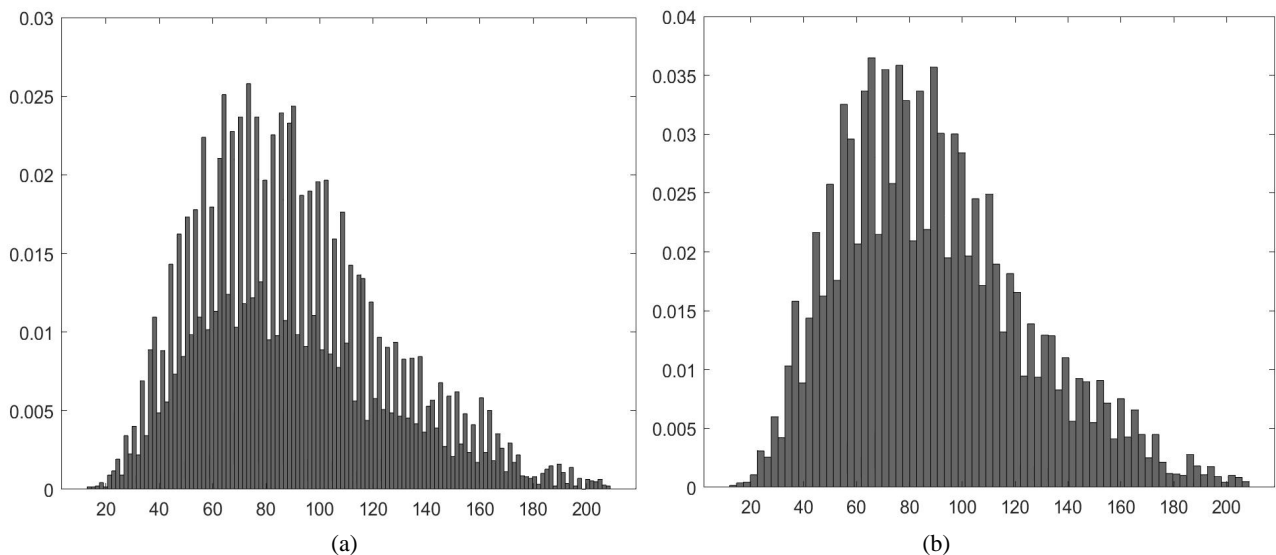
## 4. 实验结果和分析

### 研究材料的介绍以及数据准备



**Figure 1.** Example of ultrasound images (a) Original image 1; (b) Region of interest 1; (c) Original image 2; (d) Region of interest 2

**图 1.** 超声图像例图(a) 原始图像 1; (b) 有关区域 1; (c) 原始图像 2; (d) 有关区域 2



**Figure 2.** Histogram with Spikes (a) Histogram with 2 spacing; (b) Histogram with 4 spacing

**图 2.** 带有锐刺的直方图(a) 间距为 2 的直方图; (b) 间距为 4 的直方图

超声内镜中的超声探头被安置在内镜前端，在进入消化道时进行实时扫描。由于消化粘膜下肿瘤和正常消化粘膜组织的结构复杂，并且有大量随机分布的散射粒子的存在，导致超声波在传播过程中发生干涉现象，因此在超声图像中形成了亮暗不定的斑点[15]。本文采用上海市第六人民医院所提供的消化粘膜下肿瘤超声图片进行了实验，图片格式为 JPG，如图 1 所示，其中红色区域为肿瘤组织，白色区域为正常组织。我们从上海市第六人民医院共得到消化粘膜下超声图像 197 张，但由于一些图像质量较低，不能分辨消化粘膜下肿瘤组织，这些图像并未被采用，最后有 169 张肿瘤图像被用来分析消化粘膜下肿瘤的斑点的分布，有 185 张消化粘膜下正常组织图像用于分析正常组织的斑点的分布。所有超声图像中，正常组织和肿瘤组织被人工勾画出来，并对这些区域进行了灰度值采样。

本文在将图像超声斑点灰度值统计为直方图时，发现直方图中会出现如图 2 所示具有显著的锐刺，它们似乎在直方图中大致周期性地出现，每个峰值之后都是一个空的或低计数的灰度值。直方图中出现的这类锐刺(Spikes)的原因如下：锐刺是由于为了减少机器显示缓冲区的带宽，内部图像压缩和解压缩的结果；锐刺的出现是由于 A/D 转换结果较差的原因，锐刺是机器中散斑一致算法作用的结果[16]。

通过实验发现在补偿锐刺伪影的影响和保留直方图的整体形状之间，间距为 16 时是一个很好的折衷。因此在处理超声图像斑点噪声像素块直方图的时候，都固定直方图的横轴灰度级为 16。下面我们将从小超声斑点的直方图出发，探究消化粘膜下正常组织灰阶超声图像中的斑点和肿瘤灰阶超声图像中斑点分布特性。首先，对于消化道粘膜下的正常组织和肿瘤组织的灰度数据进行统计。得到的所有灰度值分为 16 组，每组 16 个灰度值。针对每组灰度值，记录其在总灰度值中的比例并绘制直方图。然后，将这 16 个灰度数据代入五种潜在的概率分布中，通过极大似然估计计算出每个概率模型的参数。最后计算并检验这五种拟合后的概率模型的拟合程度。其结果如图 3 所示。

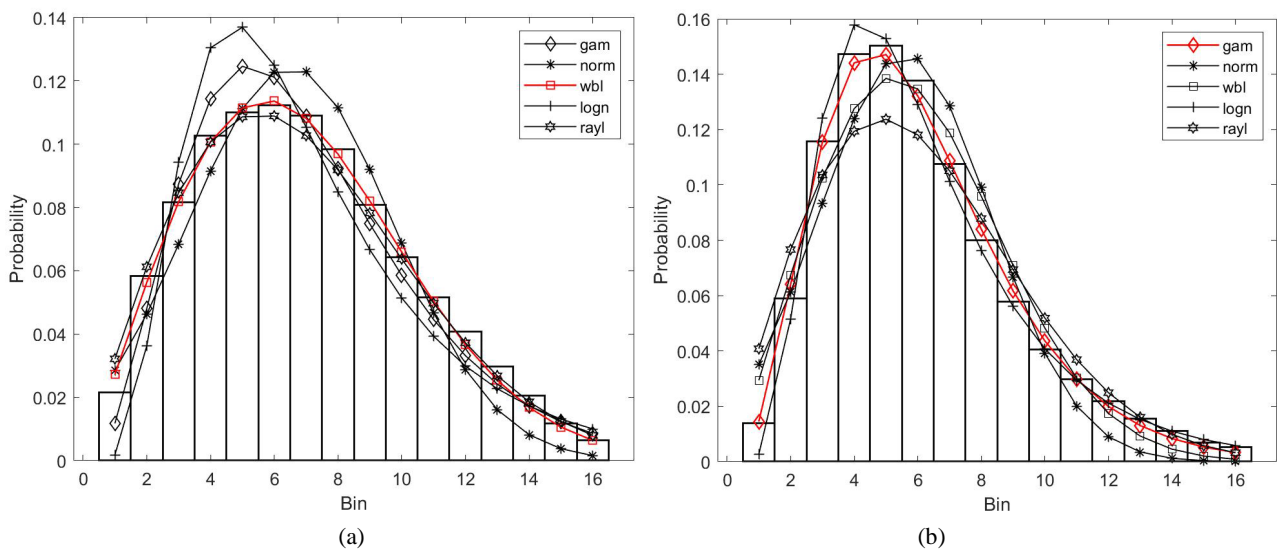


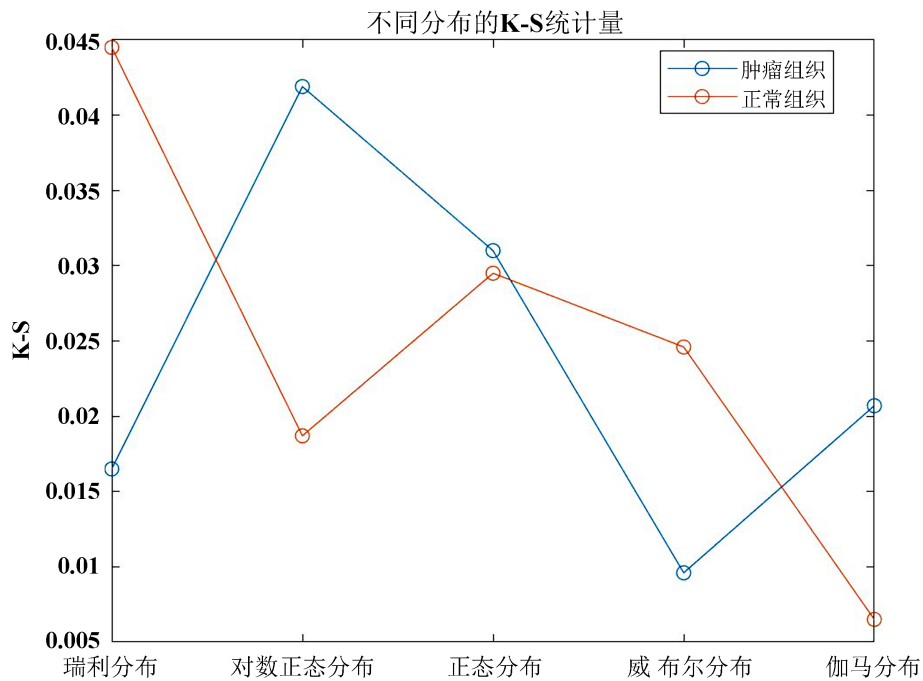
Figure 3. Histogram of fitted ultrasound images of digestive mucosa (a) Histogram of tumor tissue; (b) Histogram of normal tissue  
图 3. 拟合的消化粘膜超声图像直方图(a) 肿瘤组织直方图; (b) 正常组织直方图

表 3 给出了不同分布类型拟合正常组织和肿瘤组织的 K-S 统计量的结果。根据图 4 和表 3 数据可以看出，在正常组织和肿瘤组织的对比中，K-S 统计量显示出明显的差异，针对肿瘤组织，威布尔分布的 K-S 统计量最小，为 0.0096，明显低于伽马分布、瑞利分布、对数正态分布和正态分布。这说明威布尔分布在描述消化粘膜下肿瘤组织斑点分布时表现比伽马分布、瑞利分布、对数正态分布和正态分布更好。针对正常组织，伽马分布的 K-S 统计量最小，为 0.0065，明显低于瑞利分布、对数正态分布、威布尔分

布和正态分布的 K-S 统计量。这说明伽马分布在描述消化粘膜下正常组织斑点分布时表现比瑞利分布、对数正态分布、威布尔分布和正态分布效果更好。

**Table 3. K-S Tests**  
**表 3. K-S 检验**

	瑞利分布	对数正态分布	正态分布	威布尔分布	伽马分布
肿瘤组织	0.0165	0.0419	0.0310	0.0096	0.0207
正常组织	0.0445	0.0187	0.0295	0.0246	0.0065



**Figure 4. K-S statistics for different distributions**  
**图 4. 不同分布的 K-S 统计量**

表 4 给出了不同分布类型拟合正常组织和肿瘤组织的  $R^2$  的结果。由图 5 和表 4 数据可以看出，在正常组织和肿瘤组织的对比中， $R^2$  显示出明显的差异，针对肿瘤组织，威布尔分布的  $R^2$  最大，为 0.9956，明显高于伽马分布、瑞利分布、对数正态分布和正态分布的  $R^2$ 。这说明威布尔分布在拟合消化粘膜下肿瘤组织斑点分布时表现比伽马分布、瑞利分布、对数正态分布和正态分布更好。针对正常组织，伽马分布的  $R^2$  最大，为 0.9967，略高于对数正态分布的 0.9898，并且明显高于瑞利分布、威布尔分布和正态分布的  $R^2$ 。这说明伽马分布在拟合消化粘膜下正常组织斑点分布时表现比瑞利分布、威布尔分布、对数正态分布和正态分布更好。

**Table 4. R-square test**  
**表 4. R 方检验**

	瑞利分布	对数正态分布	正态分布	威布尔分布	伽马分布
肿瘤组织	0.8747	0.9109	0.9566	0.9956	0.9749
正常组织	0.8072	0.9898	0.9332	0.9586	0.9967

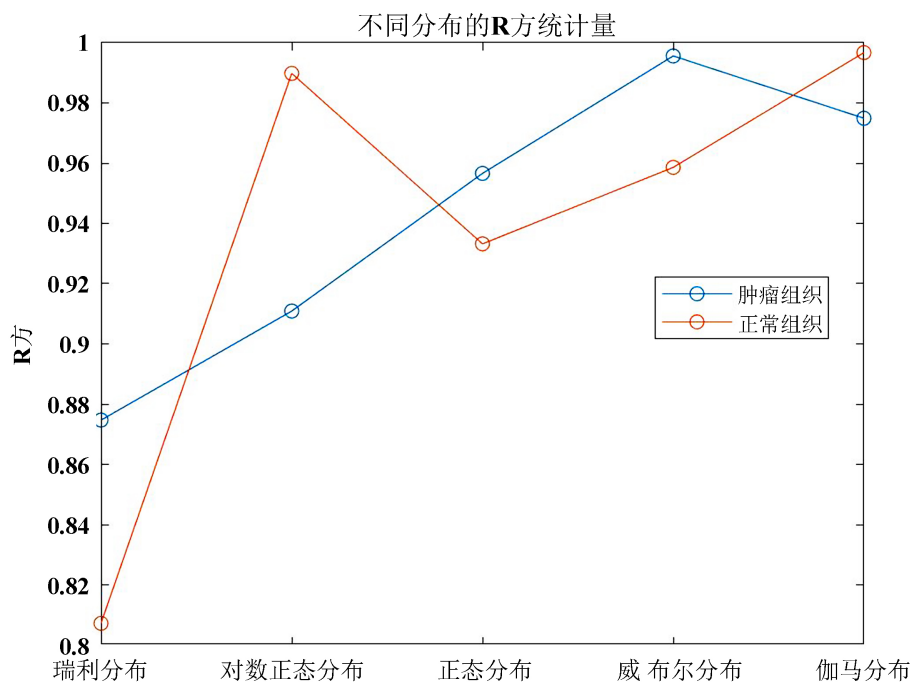


Figure 5. r-squared for different distributions

图 5. 不同分布的  $R^2$  统计量

综上所述，利用 K-S 统计量和  $R^2$  的评价指标，根据表 3 和表 4 数据发现威布尔分布在描述肿瘤组织的斑点分布时表现更好，而伽马分布在描述正常组织的斑点分布时表现更好，进一步显示了不同组织的斑点分布存在明显差异。

## 5. 结论

由于医学超声图像中斑点噪声分布依赖于底层组织，因此利用斑点分布对组织进行区分成为可能。本文利用上海市第六人民医院提供的消化道粘膜下肿瘤超声图片，通过对消化道粘膜下正常组织和肿瘤组织中斑点的概率分布模型进行分析，实验结果表明在消化粘膜正常组织的超声图像中，伽马分布表现出更好的拟合性能，其 K-S 统计量最小，仅为 0.0065，其  $R^2$  最大，为 0.9967。而在消化粘膜肿瘤组织的超声图像中，威布尔分布更为适用，其 K-S 统计量最小，为 0.0096， $R^2$  最大，为 0.9956。这也显示了不同组织类型在超声图像中的斑点分布特征存在差异，这对于超声图像的组织分割和分类具有重要意义。未来的研究可以进一步探索如何应用这些概率分布模型来增强医学影像分析的质量，以更好地满足患者的医疗需求。

## 参考文献

- [1] 杨琴, 冶丽娜, 杨红伟. 超声内镜在上消化道黏膜下病变诊治中的价值分析[J]. 中国实用医药, 2021, 16(30): 83-85.
- [2] 胡蝶. 医学超声图像的噪声分析及去噪方法研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2018.
- [3] Oelze, M.L. and Mamou, J. (2016) Review of Quantitative Ultrasound: Envelope Statistics and Backscatter Coefficient Imaging and Contributions to Diagnostic Ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, **63**, 336-351. <https://doi.org/10.1109/TUFFC.2015.2513958>
- [4] Zhang, J., Wang, C. and Cheng, Y. (2015) Comparison of Despeckle Filters for Breast Ultrasound Images. *Circuits Systems & Signal Processing*, **34**, 185-208. <https://doi.org/10.1007/s00034-014-9829-y>



- 
- [5] 边海丽. 医学超声斑点分布分析及去噪[D]: [硕士学位论文]. 杭州: 浙江大学, 2013.
- [6] Tsui, P.H, Ho, M.C. and Tai, D.I. (2016) Acoustic Structure Quantification by Using Ultrasound Nakagami Imaging for Assessing Liver Fibrosis. *Scientific Reports*, **6**, 1-9. <https://doi.org/10.1038/srep33075>
- [7] Wan, Y.L., Tai, D.I. and Ma, H.Y. (2015) Effects of Fatty Infiltration in Human Livers on the Backscattered Statistics of Ultrasound Imaging. *The Proceedings of the Institution of Mechanical Engineers*, **229**, 419-28. <https://doi.org/10.1177/0954411915585864>
- [8] Nasr, R., Falou, O. and Shahin, A. (2018) K-Nearest Neighbor Classification for the Differentiation Between Freshly Excised and Decellularized Rat Kidneys Using Envelope Statistics. 2018 *IEEE 4th Middle East Conference on Bio-medical Engineering (MECBME)*, Tunis, 28-30 March 2018, 214-218. <https://doi.org/10.1109/MECBME.2018.8402436>
- [9] Wallot, S. and Leonardi, G. (2018) Deriving Inferential Statistics from Recurrence Plots, A Recurrence-Based Test of Differences between Sample Distributions and Its Comparison to the Two-Sample Kolmogorov-Smirnov Test. *Chaos*, **28**, 085712. <https://doi.org/10.1063/1.5024915>
- [10] Ke, X., Wang, S. and Zhou, M. (2023) New Approaches on Parameter Estimation of the Gamma Distribution. *Mathematics*, **11**, 927. <https://doi.org/10.3390/math11040927>
- [11] Wang, S. and Gui, W. (2020) Corrected Maximum Likelihood Estimations of the Lognormal Distribution Parameters. *Symmetry*, **12**, 968. <https://doi.org/10.3390/sym12060968>
- [12] Ahmad, A.H. and Omar, M.E. (2021) A New one-Term Approximation to the Standard Normal Distribution. *Pakistan Journal of Statistics and Operation Research*, **17**, 381-385. <https://doi.org/10.18187/pjsor.v17i2.3556>
- [13] Tanja, V.H. (2018) Fuzzy Parameter Estimation of the Rayleigh Distribution. *Journal of Statistics and Management Systems*, **21**, 1391-1400. <https://doi.org/10.1080/09720510.2018.1519162>
- [14] Clement, N.L. and Lasky, R.C. (2020) Weibull Distribution and Analysis: 2019. 2020 *Pan Pacific Microelectronics Symposium*, HI, 10-13 February 2020, 1-5. <https://doi.org/10.23919/PanPacific48324.2020.9059313>
- [15] Rohit, S., Ricky, H. and Cailin, R. (2022) The Kidneys Are Not All Normal, Investigating the Speckle Distributions of Transplanted Kidneys. *Ultrasound in Medicine and Biology*, **49**, 1268-1274. <https://doi.org/10.1016/j.ultrasmedbio.2023.01.013>
- [16] Tao, Z., Tagare, H.D. and Beaty, J.D. (2006) Evaluation of Four Probability Distribution Models for Speckle in Clinical Cardiac Ultrasound Images. *IEEE Transactions on Medical Imaging*, **25**, 1483-1491. <https://doi.org/10.1109/TMI.2006.881376>