图像与视频扩散模型内生水印：高斯噪声分布调制与时空置乱机制研究报告

生成式视频扩散水印的演进与技术范式对比

随着生成式人工智能的爆发式增长，尤其是文本到图像（T2I）和文本到视频（T2V）扩散模型的高速迭代，合成媒体内容的版权归属验证、源头追踪和完整性保护已成为数字安全领域的紧迫课题[1]。传统的视频水印技术主要是在图像或视频渲染完成后，通过空间域像素微调或利用离散余弦变换（DCT）、离散小波变换（DWT）等经典频域变换修改系数来嵌入信息[4]。此类后处理（Post-hoc）方法虽然部署难度低，但在面对扩散模型特有的“图像重生成攻击”或“随机扩散重采样攻击”（如利用无水印扩散模型对带水印内容进行去噪重建）时，其像素级或浅层频域的水印信号极易被彻底抹除，去除率可高达 $95\% - 100\%$ [7]。此外，后处理水印往往难以在鲁棒性与感知视觉质量（即保真度）之间取得良好平衡，微小的像素修改在多帧连续播放时容易被时间注意力机制放大，从而引发画面闪烁或偏色[7]。

为了从根本上克服后处理方案的脆弱性，学术界与工业界转向了内生水印（In-Generation / Generative Watermarking）范式[11]。内生水印并不对最终的渲染像素进行显式修改，而是将水印信息的嵌入过程深度融合于扩散模型的确定性去噪轨迹之中[11]。其基本原理是利用去噪扩散隐式模型（DDIM）等采样算法所具备的数学确定性与可逆性[15]。在内容生成阶段，算法通过调制初始伪高斯噪声（Initial Latent Noise）的空间或频域结构，将水印比特预先埋入初始状态 Z_T [10]。随着反向去噪演化，扩散模型的去噪网络（如 U-Net 或双向 Transformer）在文本提示词或条件图像的引导下，自然地将初始噪声中的微小扰动编织（Weave）为图像或视频的高维时空语义纹理，使水印与生成的画面内容融为一体[5]。在水印提取阶段，检测器只需利用 VAE 编码器将受测图像或视频映射回潜空间，通过 DDIM 逆向反转（DDIM Inversion）沿着确定性的常微分方程（ODE）轨迹重构出其初始噪声 $\hat{Z}_T$，再通过特定的解调算子即可解码出原始水印[10]。这种技术被称为“噪声即水印”（Noise-as-Watermark, NaW），它提供了极强的抗编辑和抗重采样鲁棒性，且理论上对生成画质无任何损伤[13]。

基于高斯噪声分布调制的SOTA内生水印算法原理解析

傅里叶频域环形频谱对称图案调制

傅里叶频域高斯噪声调制技术以年轮水印（Tree-Ring Watermarking）为代表[12]。其核心逻辑是利用二维离散傅里叶变换（2D-DFT）将初始高斯噪声潜码 ϵ_T 从空间域映射到频域，即 $e = \mathcal{F}(\epsilon_T)$ [13]。由于自然图像的特征在频域中通常服从中心对称分布，Tree-Ring 在选定的低频或中频圆环区域 M 内，将傅里叶系数替换为预定义的、具有特定对称性的环状密钥序列 k^* [12]：

$$e[M] = k^*$$

由于频域密钥在圆环上对称分布，这种调制方式对图像的旋转、平移、裁剪和缩放（RST 变换）具有天然的数学不变性 [13]。通过逆离散傅里叶变换（2D-IDFT）将 e 还原为空间域初始高斯噪声并输入扩散模型，去噪网络会将该频域环形印记转化为图像中的隐式结构 [13]。提取水印时，系统对 DDIM Inversion 恢复的初始噪声重构值 $\hat{\epsilon}_T$ 进行傅里叶变换，并计算其在遮罩区域 M 与已知密钥 k^* 之间的检测距离 η [13]：

$$\eta = d_{\text{detection}} = 1 - \frac{1}{|M|} \sum_{i \in M} \frac{\hat{e}_i \cdot k_i^*}{|k_i^*|^2}$$

当 η 小于设定阈值时，利用非中心卡方分布（non-central χ^2 distribution）的累积分布函数 Φ_{χ^2} 评估其统计显著性（p值），即可判定水印是否存在 [13]。

在视频生成领域，直接在每一帧的初始噪声上独立套用 Tree-Ring 调制会导致严重的“时域自相关泄露” [7]。为了降低这种多帧独立注入带来的累计分布偏移，动态年轮水印（DTR / RINGet）框架提出了一种离散段分割策略 [21]。DTR 将一个长主密钥 K 离散切分为 F 个互不重叠的频域分段 $[s_1, s_2, \ldots, s_F]$，并将每个分段分别嵌入到不同帧的初始潜在变量中 [21]。这种稀疏且动态的分段嵌入极大地缓解了单一帧上由于过度频域调制引起的图像质量退化（降低了对 FID 和 FVD 指标的负面影响），在 H.264 和 H.265 视频压缩环境下实现了超过 70% 的提取鲁棒性提升 [21]。

条件截断高斯采样与双重篡改定位

VideoShield 提出了一种无需训练的条件截断高斯采样调制技术，旨在通过初始噪声的符号概率分布同时实现鲁棒的水印提取和超细粒度的时空篡改定位 [10]。

其技术过程为：首先，将原始版权水印信息比特序列 $m \in \{0,1\}^L$ 通过对称复制或密集映射进行空间维度扩展，得到大小为 $C \times H \times W$ 的特征矩阵 $\hat{m}$ [26]。为确保密码学安全性并消除规律性几何图案，VideoShield 使用安全种子 s 运行 ChaCha20 流密码算法，对 $\hat{m}$ 进行异或混淆和伪随机扩散，生成空间和通道对齐的“模板比特”矩阵 $T_P \in \{0,1\}^{F \times C \times H \times W}$ [27]。

在初始化高斯噪声潜码 Z_T 时，系统根据 T_P 中每个元素的比特值，执行非对称的高斯正负半轴条件截断采样 [27]。设 $g \sim \mathcal{N}(0, \mathbf{I})$ 为标准高斯随机变量，若当前像素点对应的模板比特 $T_{P,i} = \lambda$（其中 $\lambda \in \{0,1\}$），则该点的初始噪声值 z_i 采样如下 [10]：

$$p(z_i \mid \lambda) = (-1)^{1-\lambda} \cdot |g|$$

即当 $\lambda = 1$ 时，噪声值被迫分布在标准正态分布的正半轴（$z_i > 0$）；当 $\lambda = 0$ 时，分布在负半轴（$z_i < 0$）[27]。由于 λ 的取值概率在宏观上严格对称（各占 50%），正半轴与负半轴的混合边际概率分布依然保持完美的无偏标准正态分布[10]：

$$\int p(\beta \mid \lambda) p(\lambda) d\lambda = \frac{1}{2}\left(p(\beta \mid \lambda = 0) + p(\beta \mid \lambda = 1)\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\beta^2}{2}}$$

该采样保持了标准高斯噪声分布，使得后续去噪生成过程能够平稳运行，不引起画质降级[10]。在提取阶段，经 DDIM Inversion 导出的重构潜在噪声被转化为符号比特矩阵 I_V（正数置为1，负数置为0），其与原始 T_P 执行逐像素异或校验，一旦视频的某一空间区域或某些时间帧遭遇了外部篡改（例如画中画粘贴或剪辑），对应时空坐标下的可逆采样轨迹即告破裂，异或校验错误率将飙升至接近理论上限 0.5，从而为时空双重篡改定位提供了脆弱性度量[10]。

符号调制与伪随机纠错码（PRC）机制

为彻底杜绝 VideoShield 在大规模部署中由于固定比特图案采样而潜在带来的低频空间纹理伪影，VideoMark 引入了伪随机纠错码（PRC）符号调制采样算法[7]。其数学基础建立在带噪声学习宇称（LPN）问题的密码学硬度假设之上，包含三个核心算法算子：密钥生成 $\mathrm{KeyGen}(n, m, \mathrm{fpr}, t) \rightarrow key$，编码 $\mathrm{Encode}(key, \mathbf{m}) \rightarrow \mathbf{c} \in \{-1, 1\}^n$ 以及解码 $\mathrm{Decode}(key, \mathbf{s}) \rightarrow \mathbf{m}$[7]。

在调制阶段，VideoMark 采用符号级重写策略[17]。对于视频的第 i 帧，首先独立采集一个不带任何水印信息的纯净标准高斯噪声向量 $\epsilon_i \sim \mathcal{N}(0, \mathbf{I})$[17]。同时，该帧分配到的水印消息子段 m_i 被 PRC 编码器转换为高维伪随机符号矢量 $c_i \in \{-1, 1\}^n$[17]。最终输入扩散模型的调制初始噪声 $\hat{\epsilon}_i$ 被定义为[17]：

$$\hat{\epsilon}_i = c_i \cdot |\epsilon_i|$$

其中 $|\epsilon_i|$ 为标准高斯噪声向量的逐元素绝对值，而纠错码 c_i 仅起到控制噪声元素符号的正负调制作用[17]。因为 c_i 是纯粹的伪随机对称符号序列且与高斯幅值 $|\epsilon_i|$ 相互独立，调制后的潜在噪声 $\hat{\epsilon}_i$ 从数学期望和各阶矩上来看，依然与标准多元高斯分布完全等价[17]。相比于 VideoShield 对潜在分布的生硬截断，VideoMark 完整保留了高斯潜码的全部物理结构与边缘分布特征，在

VBench 及各项主观画质评测中展现出完全无畸变（Distortion-Free）的最高级别隐形性 [17]。

视频水印中的密钥置乱过程及其应用原理

密钥置乱打破“时空自相关”与“去同步化”两难困境

在视频扩散模型内生水印的设计体系中，安全算法学者长期受到一个固有技术瓶颈的制约，即时空自相关性与抗去同步化攻击之间的两难困境 [5]。视频扩散模型的本质是通过跨帧的时间注意力机制（Temporal Attention）在时域上对高斯潜在向量进行平滑和对齐，以生成具有强物理一致性的连续动作 [5]。如果设计者采用最直观的“一钥多帧”（即在视频所有帧的初始噪声中完全复制嵌入同一组水印图案）来追求极高的时间维度冗余性，多帧之间高度自相关的静态水印噪声图案就会在长距离时间注意力ブロック中被多轮放大，最终在生成的视频画面中投射出无法消除的周期性条纹、鬼影或网格伪影，从而造成“维数塌陷与泄露” [7]。相反，如果采用“帧帧异钥”（为每一帧分配完全随机、互不相关的独立水印密钥序列）来保障画质，时域自相关泄露虽被消除，但系统将失去对任何时间维去同步攻击（如丢帧、帧乱序、帧插值或剪辑）的防御能力 [5]。一旦帧索引对应的绝对排布发生哪怕单帧的微调偏移，反向 DDIM Inversion 提取出的噪声序列将与解调密钥完全错位，导致检测直接失败 [5]。

密钥置乱（Key Scrambling）技术的引入，正是为了在潜空间噪声调制中优雅地切断“帧绝对物理索引”与“帧加密密钥”之间的硬性绑定，从而打破上述两难困境 [5]。其核心应用原理是：通过对单一高确定性的主版权水印序列在时空维度上施加伪随机变换（置换、滑动、相位偏移等置乱算子），产生一组统计特征一致、但空间局部排列高度去相关的帧级密钥分量，将“同步敏感的序列精确解调”转换为“集合级相关性统计”或“差分序列相似度匹配” [5]。

SKeDA 中的 Shuffle-Key（SKe）置乱采样算法原理

SKeDA 提出的置乱密钥分布保持采样（SKe）模块是解决该困境的 SOTA 方案 [5]。其置乱过程的数学原理与技术细节如下：

首先，系统在生成端利用全局主密钥确定性地生成一个高维的主伪随机二进制序列作为基础密钥 [5]：

$$K_{base} \in \{-1, 1\}^N$$

为了给每个视频帧注入独立且无画面损伤的水印，SKeDA 引入了对称群 S_N 内的置换算子（Permutation Operators） [5]。定义一族与帧序号相关的伪随机置换映射关系 $\{\pi_1, \pi_2, \ldots, \pi_F\}$，每一帧的加密符号序列 K_i 被规定为基础密钥 K_{base} 在该特定置换算子下的重排结果 [5]：

$$K_i = \pi_i(K_{base}) = [K_{base}[\pi_i(1)], K_{base}[\pi_i(2)], \ldots, K_{base}[\pi_i(N)]]$$

因为置换算子只改变序列中元素的索引位置，而不改变元素值的多重集合属性，所以 K_i 中正负号的数量比例与 K_{base} 完全一致，其均值和方差保持不变 [5]。将 K_i 分别用于各帧初始噪声的

符号调制（$\hat{\epsilon}_i = K_i \odot |\epsilon_i|$），在宏观上既消除了静态高频自相关的叠加泄露（由于每一帧的水印空间几何位置被伪随机打乱），又保持了边缘概率分布的标准高斯特性[5]。
在提取阶段，SKe 置换算子的妙处在于其可逆性[5]。检测器利用 DDIM Inversion 获得各受测帧的重构初始噪声，通过符号函数提取出帧级符号序列 $\hat{c}_i \in \{-1, 1\}^N$[5]。此时，检测器无需知道这些受测帧在原始视频中的绝对帧序号，而只需对提取出的符号序列施加对应的逆置换算子 π_i^{-1}，便可将它们统一映射回基础密钥 K_{base} 的表征空间中[5]：

$$\hat{K}_{base}^{(i)} = \pi_i^{-1}(\hat{c}_i)$$

对于没有遭受篡改且包含水印的合法视频帧，其逆置换结果 $\hat{K}_{base}^{(i)}$ 将会与真实的 K_{base} 呈现极高的余弦相似度或皮尔逊相关系数；而对于恶意插入的非水印帧或遭受深度修改的区域，逆置换特征将完全退化为无意义的噪声[5]。因此，检测算法可将传统的时间序列解码转换为非同步依赖的“集合级聚合”（Set-Level Aggregation）[5]：

$$\bar{K}_{base} = \frac{1}{|S_{\text{valid}}|} \sum_{i \in S_{\text{valid}}} \pi_i^{-1}(\hat{c}_i)$$

通过在特征群组上执行均值叠加，有效抑制了信道白噪声，即使经历严重的丢帧或帧乱序攻击，只要集合中存在部分合法的逆置换分量，水印基础信号即可被高精度地提取恢复[5]。

VideoMark 中的滑动窗口偏移置乱原理

为了在处理变长视频（Variable-length Videos）生成时依然能够防御时域攻击，VideoMark 采用了一种基于一维超长母版序列的“滑动窗口偏移置乱”技术[15]。
首先，算法生成一个长度远超当前生成任务所需帧数的母版二进制随机水印序列 $M_{ext} = [m_1, m_2, \ldots, m_L]$（其中 $L \gg F$）[15]。在执行每次视频生成任务时，算法通过在区间 $[0, L - F]$ 内随机采集一个整数偏移量 p 作为置乱偏移量，以此控制时域滑窗的起点[15]。第 i 帧视频分配到的具体水印信息片段被定义为[17]：

$$m_{frame,i} = M_{ext}[p + i]$$

通过在每次视频创建中随机变换起始偏移量 p，使得即使针对相同的底层文本提示词或版权内容，由于滑窗相位的随机漂移，不同批次生成的潜在高斯空间初始化状态也呈现出彻底的独立性[15]。提取阶段，当时域丢帧导致绝对索引对齐失效时，滑动窗口偏移置乱将这一问题转化为一个经典的局部字符串近似比对问题，通过在母版序列上计算编辑距离，精准定位实际采样区间并进行对齐纠错[7]。

传统空间置乱与扩散轨迹潜空间置乱的技术特征对比

为了更直观地展现扩散轨迹内生置乱技术的先进性，下表对传统空间置乱方案与扩散潜空间内生置乱机制进行了多维度系统性对比：

置乱算法名称	作用空间与维度	数学控制逻辑	扩散模型生成质量影响	对时间去同步攻击（如丢帧）的抵御能力	计算与存储开销
Arnold猫映射 [6]	空间/变换域二维坐标（2D 空间）	二阶线性同余矩阵坐标映射变换	中等：不影响初始分布，但硬置乱与扩散动力学脱节，易产生频域杂影	极差：无法应对时间维度的对齐失效	极低（仅需矩阵坐标乘法）[33]
Josephus循环置乱 [35]	空间比特平面/像素块内（3D 分块）	基于约瑟夫斯环的变步长、变方向循环置乱	中等：像素自相关被打破，但缺乏对生成器潜码先验的自适应保护	极差：时域连续性丢失后，逆置乱完全失效	低（高维循环迭代耗时）
DNA编码交叉置乱 [36]	空间频域杂合（2D 变换域）	DNA 碱基互补配对与多维混沌序列异或	较差：强制改变频域系数能量，易导致生成画质劣化和 FID 升高	较差：仅限于单帧鲁棒性，无法处理多帧序列去同步	中等（编解码映射复杂）
SKe置换群置乱 [5]	潜空间高斯符号（1D 对称群 S_N）	基于对称置换群 π_i 的符号空间重排	无影响（优秀）：严格保持高斯边缘分布，彻底消除时空注意力纹理伪影	极强：通过逆置换集合聚合，完全免除帧对齐依赖	极低（仅执行一维索引重排）

滑动窗口相位偏移 [15]	时域切片滑动（1D 时序维）	超长母版序列下的滑窗起点随机偏移 p	无影响（优秀）：纯随机高斯符号重写，无任何几何周期模式	极强：结合编辑距离计算，可承受极高比例时域丢帧	低（主要开销为编辑距离比对）

时空一致性保护与抗视频压缩鲁棒性优化机制

级联时空细化与时序匹配模块的恢复机制

在数字视频分发网络中，水印不仅要承受常见的噪声、模糊等单帧空间域降质，还必须承受更为严苛的时域操作和恶意篡改 [7]。为此，SOTA 水印提取系统引入了两类专用的恢复与校正技术：

- 级联时空细化（**Hierarchical Spatial-Temporal Refinement, HSTR**）：该模块针对 VideoShield 等具有篡改定位功能的系统进行了深度优化 [10]。由于视频中的篡改操作通常具有高度的时空局部性（如在第 10 帧到 20 帧的左上角粘贴恶意内容），HSTR 首先通过时间细化器评估多帧初始噪声符号的对齐度，恢复遭到剪辑和重排的视频时序 [10]。随后，空间细化器利用多尺度卷积神经网络（如 UNet 骨干）捕获被局部篡改所破坏的确定性高斯分布概率断层，自动过滤由自然运动引起的像素漂移误差，输出极高信噪比的空间篡改遮罩（Spatio-Temporal Tampering Mask），实现了像素级准确度的篡改定位 [26]。
- 时序匹配模块（**Temporal Matching Module, TMM**）：主要应用于 VideoMark 等基于纠错码的无盲水印系统中 [7]。TMM 直接将时域攻击抽象为信道传输过程中的字符“擦除与替换”问题 [15]。通过将所有检测帧独立解调出的比特序列拼接成一串受损字符，计算该字符与已知全局母版序列在所有可能滑窗偏移位置下的 Levenshtein 编辑距离 [7]。即使攻击者在视频中删除了超过 50% 的中间帧，TMM 依然能利用动态规划比对在数十毫秒内恢复出最佳对齐点，重建出极为稳健的版权指纹 [7]。

SKeDA 差分注意力机制对帧间压缩失真的补偿

现代视频流媒体在分发和传输过程中，几乎均会经过 H.264（AVC）或 H.265（HEVC）等视频编码器的高倍率有损压缩 [5]。这些编解码算法的核心思想是利用时域冗余，通过帧间预测与运动补偿机制（利用 I 帧预测 P 帧与 B 帧）仅传输帧间高频残差（Residuals），同时对预测误差进行剧烈的量化和高频滤波，以压榨带宽 [38]。由于扩散模型在去噪过程中高度依赖连续帧之间的微小变化，这种对残差的过滤会导致通过 DDIM Inversion 逆向重建初始噪声时产生极大的累积漂移误差，使得 P 帧和 B 帧对应的重构噪声高频分量出现大面积退化 [2]。

SKeDA 提出的差分注意力（Differential Attention, DA）机制专门针对此类非线性压缩量化失真进行了补偿优化 [5]。在去噪迭代的逆向反转过程中，DA 并不平权地分配时间步注意力权重，而是首先通过计算相邻帧特征的绝对变化率，显式地构建时域潜在差分图 [5]：

$$\Delta Z_t^{(i)} = \left| Z_t^{(i)} - Z_t^{(i-1)} \right|$$

通过该潜在差分算子，DA 能够高灵敏度地识别出视频序列中哪些时空区域处于由于物体运动剧烈而导致运动补偿编码误差极大的“高噪声预测区”（通常为残差项被大量丢弃的 P/B 帧部分），以及哪些区域处于运动平缓、高频空间信息被完美继承的“高感知保真区”（通常为静态背景或缓慢漂移区域）[5]。

在自注意力特征聚合时，DA 会引入一个基于时域差分分布的自适应校正矩阵 $\mathbf{W}_t^{(i)}$ 并将其注入到自注意力权重计算中 [5]：

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}} \odot \mathbf{W}_t^{(i)}\right) V$$

该校正矩阵极大地衰减了那些遭受编解码器重度量化的 P/B 帧动态区域的注意力响应，而对空间高频谱保留完整、去噪轨迹重构精度极高的关键区域分配了主导权重 [5]。这种自适应权重调配机制极大地提高了反向求导的信噪比，在不改变扩散模型原始推理逻辑的前提下，显著补偿了 H.264/H.265 量化带来的残差畸变，使最终的水印提取比特准确率在中高压缩率（如 CRF=30 甚至更低比特率）下较传统方案实现了 $5\% - 20\%$ 的绝对性能跨越 [5]。

多尺度时空注入与跨模态对齐

除了去噪轨域的显式调制外，近年来如 WaTeRFlow 和 GenPTW 等前沿多尺度嵌入框架也提供了富有启发性的优化路径 [11]。这些工作不再单方面依赖于对初始噪声的单点调制，而是探索多尺度时空特征的协同注入 [8]：

- 流引导统一合成引擎（**FUSE**）与时间一致性损失（**TCL**）：WaTeRFlow 框架通过光学流（Optical Flow）扭曲技术在编码器和解码器训练期间引入 TCL，将空间特征与连续运动流强力对齐 [41]。这使得即使视频从单帧静态图像通过 Image-to-Video（I2V）算法转换为多秒长视频，每一帧生成的水印信号仍能与光流场保持运动一致，解决了跨模态生成过程中的水印流失难题 [41]。
- 跨模态交叉注意力与空间融合（**GenPTW**）：GenPTW 框架将版权所有权追踪与篡改定位在 VAE 潜空间层面执行了联合建模 [11]。它在生成阶段通过将结构化的多比特水印信号送入跨注意力融合层，使之与文本/图像的潜在语义对齐，再利用空间融合模块引入多尺度的拉普拉斯高频敏感机制 [11]。通过这种联合两阶段嵌入，模型即便遭受极具破坏性的复合攻击（如同时经历高比例裁剪、加噪、再通过有损 H.264 编码分发），多尺度提取器仍能凭借对高低频特征的时空协同解调，稳定地输出高精度验证结果，展示出工业级应用环境下的广阔前景 [11]。

总结

图像与视频扩散模型内生水印技术，正沿着“无偏高斯潜空间调制”与“时空密钥置乱”的双轨并进路线，向着工业级的高隐蔽性、强鲁棒性以及多维度篡改可控定位方向加速演进 [10]。本报告系统梳理的 Tree-Ring 频域调制、VideoShield 条件截断采样、VideoMark 纠错符号调制等核心理论，有力证明了通过在生成源头精细干预噪声分布，能够达成生成质量与水印强度的和谐共存 [10]。密钥置乱机制作为攻克视频时域自相关泄露与时序去同步化瓶颈的关键钥匙，其通过置换群（SKeDA-SKe）及滑动偏移（VideoMark）等手段，彻底解耦了物理时序依赖，为业界提供了具备高度

普适性的、面向大规模视频分发的数字安全主动防御体系[5]。未来的研究不仅会进一步探索在纯盲检测场景下的高容量内生编码方案，还将致力于结合后量子密码学与抗共谋机制（Collusion-Resistant），以更强劲的技术手段，确保人工智能生成内容的可信、合规与源头可溯[42]。

引用的著作

1. Digital Watermarking Technology for AI-Generated Images: A Survey - ResearchGate, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/389085052_Digital_Watermarking_Technology_for_AI-Generated_Images_A_Survey
2. SKeDA: A Generative Watermarking Framework for Text-to-video Diffusion Models - arXiv, 访问时间为 五月 28, 2026, https://arxiv.org/abs/2603.00194
3. Secure and Efficient Watermarking for Latent Diffusion Models in Model Distribution Scenarios - arXiv, 访问时间为 五月 28, 2026, https://arxiv.org/html/2502.13345
4. Secure and flexible image watermarking using IWT, SVD, and chaos models for robustness and imperceptibility - PMC, 访问时间为 五月 28, 2026, https://pmc.ncbi.nlm.nih.gov/articles/PMC11871132/
5. SKeDA: A Generative Watermarking Framework for Text-to-video Diffusion Models - arXiv, 访问时间为 五月 28, 2026, https://arxiv.org/html/2603.00194v1
6. A blind scene-based watermarking for video copyright protection - ResearchGate, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/257585211_A_blind_scene-based_watermarking_for_video_copyright_protection
7. VideoMark: A Distortion-Free Robust Watermarking Framework for Video Diffusion Models, 访问时间为 五月 28, 2026, https://arxiv.org/html/2504.16359v3
8. VideoMark: A Distortion-Free Robust Watermarking Framework for Video Diffusion Models, 访问时间为 五月 28, 2026, https://www.semanticscholar.org/paper/VideoMark%3A-A-Distortion-Free-Robust-Watermarking-Hu-Li/0fd39e04edd553af0817dcab86497e1d8f25ef90
9. LVMark: Robust Watermark for Latent Video Diffusion Models | IEEE Journals & Magazine, 访问时间为 五月 28, 2026, https://ieeexplore.ieee.org/document/11495243/
10. VIDEOSHIELD: REGULATING DIFFUSION-BASED VIDEO GENERATION MODELS VIA WATERMARKING - ICLR Proceedings, 访问时间为 五月 28, 2026, https://proceedings.iclr.cc/paper_files/paper/2025/file/8227285e32f70e07fa3a247f3a48006d-Paper-Conference.pdf
11. GenPTW: Latent Image Watermarking for Provenance Tracing and Tamper Localization, 访问时间为 五月 28, 2026, https://arxiv.org/html/2504.19567v2
12. Scalable Dual Fingerprinting for Hierarchical Attribution of Text-to-Image Models - CVF Open Access, 访问时间为 五月 28, 2026, https://openaccess.thecvf.com/content/ICCV2025/papers/Fei_Scalable_Dual_Fingerprinting_for_Hierarchical_Attribution_of_Text-to-Image_Models_ICCV_2025_paper.pdf
13. Tree-Rings Watermarks: Invisible Fingerprints for Diffusion Images - NIPS, 访问时

间为 五月 28, 2026, https://proceedings.neurips.cc/paper_files/paper/2023/file/b54d1757c190ba20dbc4f9e4a2f54149-Paper-Conference.pdf
14. GenPTW: Latent Image Watermarking for Provenance Tracing and Tamper Localization, 访问时间为 五月 28, 2026, https://ojs.aaai.org/index.php/AAAI/article/view/37412/41374
15. VideoMark: A Distortion-Free Robust Watermarking Framework for Video Diffusion Models, 访问时间为 五月 28, 2026, https://arxiv.org/html/2504.16359v1
16. VideoMark: A Distortion-Free Robust Watermarking Framework for Video Diffusion Models - arXiv, 访问时间为 五月 28, 2026, https://arxiv.org/pdf/2504.16359
17. VideoMark: A Distortion-Free Robust Watermarking Frame- work for Video Diffusion Models - OpenReview, 访问时间为 五月 28, 2026, https://openreview.net/attachment?id=V1KnWicf7A&name=pdf
18. GROW: Watermark Generation with Progressive Guidance for Diffusion Models - CVF Open Access, 访问时间为 五月 28, 2026, https://openaccess.thecvf.com/content/CVPR2026/papers/Luo_GROW_Watermark_Generation_with_Progressive_Guidance_for_Diffusion_Models_CVPR_2026_paper.pdf
19. arXiv Papers of Watermarking - Hongsong Wang, 访问时间为 五月 28, 2026, https://hongsong-wang.github.io/Watermarking/
20. Tree-Ring Watermarks: Fingerprints for Diffusion Images that are Invisible and Robust, 访问时间为 五月 28, 2026, https://huggingface.co/papers/2305.20030
21. DTR: Dynamic Tree-Ring Watermarking Framework for Diffusion-Based Video Generation - IEEE Xplore, 访问时间为 五月 28, 2026, https://ieeexplore.ieee.org/iel8/10887540/10887541/10888152.pdf
22. digital watermarking techniques: Topics by Science.gov, 访问时间为 五月 28, 2026, https://www.science.gov/topicpages/d/digital+watermarking+techniques
23. DTR: Dynamic Tree-Ring Watermarking Framework for Diffusion-Based Video Generation, 访问时间为 五月 28, 2026, https://ieeexplore.ieee.org/document/10888152
24. RINGet: A Robust Watermarking Framework for Diffusion-Based Video Generation, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/398642759_RINGet_A_Robust_Watermarking_Framework_for_Diffusion-Based_Video_Generation
25. VideoShield: Regulating Diffusion-based Video Generation Models via Watermarking, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/388401797_VideoShield_Regulating_Diffusion-based_Video_Generation_Models_via_Watermarking
26. TAG-WM: Tamper-Aware Generative Image Watermarking via Diffusion Inversion Sensitivity, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/393184562_TAG-WM_Tamper-Aware_Generative_Image_Watermarking_via_Diffusion_Inversion_Sensitivity
27. [Literature Review] VideoShield: Regulating Diffusion-based Video Generation Models via Watermarking, 访问时间为 五月 28, 2026,

https://www.themoonlight.io/en/review/videoshield-regulating-diffusion-based-video-generation-models-via-watermarking
28. VideoMark: A Distortion-Free Robust Watermarking Framework for Video Diffusion Models, 访问时间为 五月 28, 2026, https://arxiv.org/html/2504.16359v2
29. [2504.16359] VideoMark: A Distortion-Free Robust Watermarking Framework for Video Diffusion Models - arXiv, 访问时间为 五月 28, 2026, https://arxiv.org/abs/2504.16359
30. VideoMark: A Distortion-Free Robust Watermarking Framework for Video Diffusion Models, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/391057871_VideoMark_A_Distortion-Free_Robust_Watermarking_Framework_for_Video_Diffusion_Models
31. SKeDA: A Generative Watermarking Framework for Text-to-video Diffusion Models - arXiv, 访问时间为 五月 28, 2026, https://arxiv.org/pdf/2603.00194
32. (PDF) An Overview of Digital Video Watermarking - ResearchGate, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/317352362_An_Overview_of_Digital_Video_Watermarking
33. An image encryption scheme using PRESENT-RC4, chaos and secure key generation, 访问时间为 五月 28, 2026, https://pmc.ncbi.nlm.nih.gov/articles/PMC12663476/
34. Robust Image Watermarking Using Bidirection-Interactive and Context-Aware Networks, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/389192217_Robust_Image_Watermarking_Using_Bidirection-Interactive_and_Context-Aware_Networks
35. Chaotic encryption algorithm with scrambling diffusion based on the Josephus cycle, 访问时间为 五月 28, 2026, https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2023.1191793/full
36. A Non-Blind Image Watermarking Method for Copyright Protection - JOURNAL OF UNIVERSITY OF BABYLON for Pure and Applied Sciences, 访问时间为 五月 28, 2026, https://journalofbabylon.com/index.php/JUBPAS/article/download/4182/3187/
37. DWT-DCT-SVD based watermarking | Request PDF - ResearchGate, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/4346308_DWT-DCT-SVD_based_watermarking
38. Design Principles for Orthogonal Moments in Video Watermarking - ResearchGate, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/391711005_Design_Principles_for_Orthogonal_Moments_in_Video_Watermarking
39. A Novel Deep Video Watermarking Framework with Enhanced Robustness to H.264/AVC Compression | Request PDF - ResearchGate, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/375031253_A_Novel_Deep_Video_Watermarking_Framework_with_Enhanced_Robustness_to_H264AVC_Compression

40. 0970-2555 Volume : 53, Issue 12, No.3, December : 2024 UGC CARE Group-1 144 HYBRID STR - Indian Institution of Industrial Engineering, 访问时间为 五月 28, 2026 ， http://www.journal-iiie-india.com/1_dec_24/19.3_dec.pdf
41. kuai-lab/cvpr26_WaTeRFlow: [CVPR2026] WaTeRFlow: Watermark Temporal Robustness via Flow Consistency Resources - GitHub, 访问时间为 五月 28, 2026, https://github.com/kuai-lab/cvpr26_WaTeRFlow
42. Deep Learning-Based Video Watermarking: A Robust Framework for Spatial–Temporal Embedding and Retrieval - MDPI, 访问时间为 五月 28, 2026, https://www.mdpi.com/1999-5903/18/2/104
43. GenPTW: In-Generation Image Watermarking for Provenance Tracing and Tamper Localization - arXiv, 访问时间为 五月 28, 2026, https://arxiv.org/html/2504.19567v1
44. Invisible Signatures: A Survey of Digital Watermarking Technology, 访问时间为 五月 28, 2026， https://yage.ai/share/dwt-watermark-survey-en-20260526.html
45. Digital Watermarking and Steganography | Request PDF - ResearchGate, 访问时间为 五月 28, 2026, https://www.researchgate.net/publication/263606329_Digital_Watermarking_and_Steganography