基于潜在空间 HSIC 正则化的图像解耦生成方法

李元昊，应方立

（华东理工大学信息科学与工程学院，上海 200237）

摘　要： 学习解耦表征以提升图像生成模型的可控性是计算机视觉领域的重要研究方向。然而，现有解耦表征学习方法存在两大局限性：一是依赖大规模标注数据；二是难以有效处理特征间的复杂依赖关系。为突破这些限制，提出一种基于希尔伯特-施密特独立性准则（HSIC）的通用解耦生成方法。该方法创新性地将 HSIC 这一非参数统计方法转化为生成模型潜在空间的独立性正则化机制，通过施加 HSIC 正则项优化非线性依赖关系的度量目标，引导模型学习独立的特征表示。具体而言，通过实验将 HSIC 融入两类主流生成模型架构的优化过程：在变分自编码器（VAE）模型类中，通过结合变分推断重构与 HSIC 正则项，优化潜在分布的解耦性；在扩散模型（DM）类中，通过将 HSIC 正则项嵌入反向过程的时间步优化，逐步实现渐进式特征解耦。实验结果表明，这种能够在不同模型架构中实现的通用方法提升了潜在表示的独立性，且在无监督场景下保持稳定性能，为建模特征间复杂依赖关系提供了新途径。为进一步验证解耦空间的语义一致性，通过潜在空间插值实验，生成轨迹更加平滑的结果，证明了 HSIC 正则化有效构建了线性可分的解耦空间。在评估体系方面，采用标准解耦指标与基于 HSIC 的自定义指标进行双重验证，结果两者呈正相关，证实了解耦评价标准的客观性。

关键词： 图像生成模型；解耦表征学习；希尔伯特-施密特独立性准则；正则化机制；潜在空间插值

中图分类号： TP391　　**文献标志码：** A　　**DOI：** 10.19678/j.issn.1000-3428.0252270

Image Disentangling Generation Method Based on Latent Space HSIC Regularization

LI Yuanhao, YING Fangli

(School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

【Abstract】 Learning disentangled representations to enhance the controllability of image generation models is a key research direction in computer vision. However, existing methods have two major limitations: reliance on large-scale annotated data and difficulty in handling complex dependencies between features. To address these issues, this study proposes a universal generative disentangling method based on the Hilbert-Schmidt Independence Criterion (HSIC). This method innovatively converts HSIC into an independence regularization mechanism for the latent space of generative models. By incorporating HSIC regularization terms, it optimizes the measurement objective of nonlinear dependency relationships and guides the model in learning independent feature representations. Specifically, the method integrates HSIC into two mainstream generative model architectures. For the Variational AutoEncoder (VAE) class, it combines variational inference with HSIC regularization to optimize latent distribution disentanglement. For the Diffusion Model (DM) class, it gradually achieves progressive feature disentangling by embedding the HSIC regularization term into the time-step optimization of the reverse process. The experimental results show that this universal method, which can be implemented in different model architectures, enhances the latent representation independence and maintains stable performance in unsupervised settings, offering a novel approach to model complex feature dependencies. To further verify the semantic consistency of the disentangling space, latent space interpolation experiments are conducted to generate smoother trajectories, demonstrating that HSIC regularization constructs a linearly separable disentangling space. Additionally, a dual validation using standard disentangling metrics and HSIC-based custom metrics is performed, which reveals a positive correlation and confirms the objectivity of the disentangling evaluation criteria.

【Key words】 image generation model; disentangled representation learning; Hilbert-Schmidt Independence Criterion (HSIC); regularization mechanism; latent space interpolation

基金项目： 山东省重点研发计划（2022SFGC0104）。

作者简介： 李元昊（CCF 学生会员），男，硕士研究生，主研方向为图像生成模型、解耦表征学习；应方立（通信作者），讲师、博士。

收稿日期： 2025-03-27　　**修回日期：** 2025-05-23　**E-mail：** yfangli@ecust.edu.cn

计算机工程 www.ecice006.com

0 引言

在计算机视觉图像生成领域，解耦表征学习是实现模型可控性的重要手段[1]。解耦表征学习旨在通过挖掘数据内部生成因子，将其分解为独立且可解释的潜在表示，实现对数据的深入理解和有效表征控制[2-3]。然而，早期基于统计独立性假设的传统解耦表征学习方法存在两大局限性：一是由于独立性假设无法保证潜在表示与可理解语义属性对齐，因此严重依赖大规模标注数据；二是由于独立性假设与真实数据间的冲突，因而难以有效建模特征间的复杂依赖关系[4]。这些问题限制了传统方法在无监督复杂图像生成任务中的应用。

目前，在图像解耦生成任务中，主流解耦表征学习方法根据生成模型架构的差异可分为三类：基于变分自编码器（VAE）的方法[5-6]、基于生成对抗网络（GAN）的方法[7]和基于扩散模型（DM）的方法[8]。这些方法在可控生成和图像翻译等下游任务中取得了显著成果[1]，其应用价值已通过多项实际场景验证，例如，图像修复[9]、水下图像增强[10]等。然而，在数据具有多维度、高耦合性且隐含非线性统计关系的复杂场景下，通常需要引入标注数据这样的监督信号以提升解耦效果，同时独立性假设强制将联合分布分解为边缘分布乘积，导致对高阶统计关系的建模能力有限[4]。具体而言，基于 VAE 的解耦表征学习方法（如 β-VAE[11]、VQ-VAE[12]等）通常通过引入正则化约束来实现解耦。然而，其生成质量受限于潜在空间的分布假设，难以捕捉复杂数据中潜在变量与语义属性的双向非线性依赖关系。基于 GAN 的方法（如 InfoGAN[7]、StyleGAN[13]等）通过潜在空间结构化分解、判别器增强等实现解耦，但其因解耦与生成目标梯度冲突，导致训练稳定性差，且高阶语义解耦被证实需要依赖大量标注数据。基于 DM 的方法（如 DisDiff[8]等）在马尔可夫扩散-去噪框架[14]中，通过编码器分离与梯度正则化实现解耦，但其过程的可解释性仍有待提升。在没有标注数据的无监督复杂图像生成场景下，这些方法在特征独立性等方面仍存在明显缺陷。

为克服这些缺陷，本文提出了一种基于希尔伯特-施密特独立性准则（HSIC）[15]的通用图像解耦生成方法，支持显式潜在变量的多模型架构。HSIC 作为一种非参数统计方法，计算过程不依赖预定义概率分布和标注数据，可直接通过样本计算独立性度量，并能有效捕捉变量间的非线性关系[15]。具体而言，本文在主流生成模型架构（如 VAE、DM 这样具有显式潜在变量的模型架构）的编码过程中引入 HSIC 正则项，从独立性度量方面学习解耦表征。由于 HSIC 通过无监督方式直接度量分布独立性，与解耦表征目标高度契合，因此该方法不仅减少了模型对标注数据的依赖，还突破传统统计独立性假设，为具有非线性、高耦合等特征的复杂数据关系建模提供了新途径。

本研究的贡献如下：

1）提出了一种跨架构解耦生成方法，通过在支持显式潜在变量的主流生成模型（如基于 VAE、DM 的方法）中引入 HSIC 正则项，提升了潜在表示的独立性，实现图像特征的解耦表征。

2）设计了一种基于 HSIC 的无监督解耦评估框架，通过样本直接计算 HSIC 值，无需人工标注，并联合标准解耦指标与 HSIC 自定义指标进行双重验证，从而提升解耦评估的可解释性。

3）通过图像重建、潜在空间插值以及消融实验，验证了所提方法的有效性，同时证明了基于 HSIC 正则化解耦的空间具有线性可分性。

1 相关工作

1.1 解耦表征学习

解耦表征学习旨在从观测数据中分离出可解释、泛化性强且可控的潜在表示[4]。现有的解耦表征方法主要基于不同的生成模型架构。基于 VAE 的方法，比如 β-VAE[11]，其通过 KL 散度约束潜在空间分布，强制潜在变量趋近独立高斯先验，从而分离不同生成因子；文献[16]提出的 JointVAE 通过融合连续和离散潜在变量增强解耦能力。这类方法基于变分推断框架，理论可解释性强，但依赖的低阶统计约束难以建模高阶语义关联，导致在需复杂因子交互的合成场景中性能受限。基于 GAN 的方法，如 InfoGAN[7]通过最大化潜在变量与生成特征的互信息实现解耦，其优势在于生成图像质量高，但因无法自适应数据真实因子结构，导致解耦能力受限于预设的潜在变量维度，且对抗训练稳定性差。基于 DM 的方法，如 DisDiff[8]和 LCG-DM[17]，借鉴了 VAE 的结构化潜在空间设计思想与 GAN 的高质量生成目标，但其依赖任务定制化设计，导致跨场景泛化性不足。例如，LCG-DM 假设潜在特征服从线性可分分布，依赖人工定义的潜在条件来引导生成过程，此类强假设在复杂场景中难以成立，导致跨场景泛化性受限。

综上所述，现有方法在解耦效果、生成质量与跨场景泛化性之间难以权衡。如基于 VAE 的方法受

限于低阶统计量约束，难以建模复杂的高阶语义依赖，导致在细粒度控制的场景中性能受限；基于 GAN 的方法因对抗训练不稳定且预设潜在维度与真实数据不匹配，导致解耦能力弱；基于 DM 的方法则需任务定制化设计，依赖先验知识，导致跨模态生成场景中泛化性不足。

1.2 HSIC

现有解耦方法通常依赖协方差或互信息等传统独立性度量约束潜在变量，但协方差无法捕捉变量间的非线性依赖关系，互信息难以准确建模高维数据或复杂分布变量间的高阶依赖关系。为此，HSIC 被引入解耦表征学习方法。HSIC 是一种基于核方法的非参数统计工具，其核心思想是通过将数据映射到再生核希尔伯特空间（RKHS），捕捉变量间的非线性与高阶依赖关系[15]。具体而言，对于随机变量 x 和 y 的联合分布 p_{xy}，HSIC 定义为通过计算 RKHS 中交叉协方差算子 $\boldsymbol{C}_{xy}$ 的平方希尔伯特-施密特（HS）范数：

$$\mathrm{HSIC}(p_{xy},\boldsymbol{F},\boldsymbol{G})=\|\boldsymbol{C}_{xy}\|^2_{\mathrm{HS}} \tag{1}$$

式中：$\boldsymbol{F}$ 和 $\boldsymbol{G}$ 为由核函数诱导的 RKHS；$\boldsymbol{C}_{xy}$ 描述了从空间 $\boldsymbol{G}$ 到空间 $\boldsymbol{F}$ 的线性映射；HS 范数的作用是衡量 $\boldsymbol{F}$ 和 $\boldsymbol{G}$ 之间 $\boldsymbol{C}_{xy}$ 的强度，其平方值通过核矩阵的迹计算。基于从 p_{xy} 中独立抽取的 m 个观测样本 $Z=\{(x_1,y_1),\cdots,(x_m,y_m)\}$，其经验估计可简化为：

$$\mathrm{HSIC}(Z,\boldsymbol{F},\boldsymbol{G})=(m-1)^{-2}\mathrm{tr}(\boldsymbol{K}\cdot\boldsymbol{H}\cdot\boldsymbol{L}\cdot\boldsymbol{H}) \tag{2}$$

式中：$\boldsymbol{K}$ 和 $\boldsymbol{L}$ 分别为输入数据 x 和 y 的核矩阵，定义为 $\boldsymbol{K}_{ij}=k(x_i,x_j)$ 和 $\boldsymbol{L}_{ij}=l(y_i,y_j)$，$k$ 和 l 是核函数；$\boldsymbol{H}$ 为中心化矩阵，定义为 $H_{ij}=\delta_{ij}-m^{-1}$，$\delta$ 为狄拉克评估算子；tr 为矩阵迹运算操作，则 $\boldsymbol{K}\cdot\boldsymbol{H}\cdot\boldsymbol{L}\cdot\boldsymbol{H}$ 乘积是计算中心化后的 x 和 y 在 RKHS 中的交叉协方差。

HSIC 作为一种严格的独立性判据，通过核函数映射来衡量两个变量间的统计依赖性。经验 HSIC 估计值恒为非负数，其数值越低表示变量间独立性越强。当且仅当 HSIC 值为 0 时，变量严格独立。与传统解耦方法不同，HSIC 能够通过核函数隐藏式捕捉非线性、非单调甚至高阶依赖关系。此外，HSIC 基于样本直接计算，无需依赖标注数据，适用于高维复杂场景下的无监督学习任务[18]。得益于其独立性度量，HSIC 已被广泛应用于多个领域。在解耦表征学习方面，HSIC-InfoGAN 通过替换使用 HSIC 近似互信息来优化 GAN 的解耦表征[19]。在持续学习方面，DualHSIC 利用 HSIC 抑制新旧任务间干扰，提升模型抗遗忘能力[20]。在评估基准方面，HSIC 凭借其无需参数假设与计算高效性成为独立性度量基准[15]。

1.3 解耦评估方法

解耦表征学习的评估需同时兼顾生成质量与解耦效果[4]。传统生成质量评估指标（如 FID、IS）虽然能衡量生成性能，但难以直接反映解耦效果，因此需依赖针对性的解耦评估度量。现有解耦评估指标可分为三类：基于干预的指标、基于预测的指标和基于信息的指标[21]。基于干预的指标是通过人为扰动潜在变量并观察生成结果的变化程度，衡量因子独立性。典型方法如 FactorVAE score。FactorVAE score 通过线性分类器判断潜在变量扰动后的生成变化，但需人工预设扰动维度，存在计算复杂度高、主观性强等问题[22]。基于预测的指标是利用标注数据建立潜在变量与已知因子的映射关系，评估解耦语义一致性，如 DCI。DCI 是通过训练回归模型量化潜在变量对标注因子的预测能力，但依赖完整标注数据，难以扩展至无监督场景[23]。基于信息的指标是直接量化潜在变量间的统计独立性，衡量解耦强度，如 TAD 评分。TAD 通过计算接收者操作特征曲线下面积（AUROC）衡量潜在变量与数据属性的独立性，量化模型解耦程度，但需属性标签的对应支持而常用于特定数据集[24]。

上述指标分别从干预可控性、语义对齐性与统计独立性的不同维度评估解耦效果，但单一指标仅能反映解耦的某一侧面，需联合使用多指标实现全面评估。此外，现有方法多数依赖大量标注数据，严重制约了其在真实无监督场景中的应用。

2 方法

在这一部分，本文将详细描述基于 HSIC 理论基础的通用解耦生成方法设计，包括在 VAE 类模型和 DM 类模型中的具体设计，以及解耦评估方法。

2.1 通用解耦生成方法设计

为了解决解耦表征学习中的问题，本文提出了一种基于 HSIC 的通用解耦方法，如图 1 所示（彩色效果见《计算机工程》官网 HTML 版，下同），其核心在于通过非参数统计方法建立潜在表示的 HSIC 独立正则化机制。

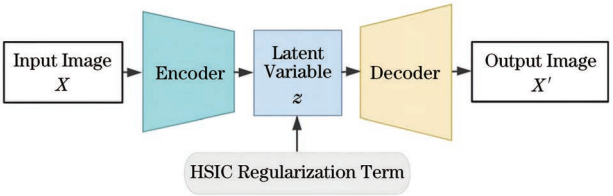


图 1 基于 HSIC 的通用解耦方法

Fig. 1 General disentanglement method based on HSIC

2.1.1 VAE 类模型中的解耦生成

在 VAE 类模型中，本方法将 HSIC 构建为潜在空间独立性正则项，并将其系统性地融入 VAE 模型的优化目标中，通过变分推断约束潜在分布。

在传统 VAE 框架中，模型是通过编码器将输入图像数据转换为潜在空间的变量表示，然后通过解码器从潜在空间的变量表示重构输入数据。根据 VAE 的变分推断框架，模型优化目标为最大化证据下界，其目标函数由两部分组成：重构损失 L_{recon} 和 KL 散度 D_{KL}。其核心是通过 KL 散度约束潜在分布逼近先验分布，但未显式建模潜在变量维度间的依赖关系。而复杂数据潜在变量维度间往往存在复杂的耦合关系，导致生成过程难以独立控制特定特征。

为此，本文引入 HSIC 作为正则项，约束潜在变量组间的独立性，以核方法隐式建模潜在变量间的高阶依赖关系，提升潜在变量的维度独立性。其中，HSIC 正则项定义：将潜在变量 z 划分为 k 个因子组$\{z_1, z_2, \cdots, z_k\}$，每组对应一个语义因子（如颜色、形状）。对每个因子组 z_i，采用高斯核函数来计算核矩阵 $\boldsymbol{K}_i$，然后对任意两个因子组 z_i 和 z_j 根据式(2)计算其 HSIC 值，最后总正则项值为所有因子组对的 HSIC 之和。

$$L_{\text{HSIC}}=\frac{1}{k(k-1)}\sum_{i\neq j}\text{HSIC}(z_i,z_j) \tag{3}$$

根据该正则结果，在模型推断过程中，通过最小化这个总正则项结果，模型会强制不同因子组的潜在变量满足统计独立性，从而实现解耦表征。

在 VAE 模型中引入 HSIC 正则项，等价于在其证据下界中引入拉格朗日乘子约束，即 HSIC-VAE 模型的总损失为：

$$L_{\text{VAE}}=L_{\text{recon}}+\beta D_{\text{KL}}+\lambda L_{\text{HSIC}} \tag{4}$$

式中：β 和 λ 分别为控制 KL 散度的权重和 HSIC 项的调节系数，且 VAE 模型中 $\beta=1$。通过调节 λ，可控制模型正则化力度，进而控制解耦强度。当 λ 增大时，模型强制因子组间独立性增强，实现属性解耦表征。在对其他方面的影响上，假设真实数据在生成过程满足潜在因子组独立，根据变分推断理论，HSIC-VAE 模型的证据下界等价于在 L_{HSIC} 和 D_{KL} 足够小的约束条件下最大化似然 L_{recon}。通过拉格朗日对偶性，正则项系数 λ 对应约束边界的松弛变量，其调节等价于调整解耦强度与重构精度间的权衡。因此增大 λ 虽然增强解耦效果，但会使生成质量降低。

同理，也可将 HSIC 正则项引入到 JointVAE 模型中，用来在拆开连续和离散变量基础上增强连续变量内部间的独立性约束。基于不同模型的解耦改进如图 2 所示。如图 2(a)所示，对于联合建模连续变量 z 和离散变量 c 的 JointVAE 框架，主要针对连续变量进行额外的 HSIC 正则。如式(3)所示，对连续因子组间进行 HSIC 值计算，得到总 HSIC 正则项 L_{HSIC}，因此在该基线模型上施加了 HSIC 正则项的 HSIC-JointVAE 模型的总损失为：

$$\begin{aligned}L_{\text{JointVAE}}&=L_{\text{recon}}+\beta_{\text{cont}}L_{\text{cont}}+\beta_{\text{disc}}L_{\text{disc}}=\\&L_{\text{recon}}+\beta_{\text{cont}}\left|C_{\text{cont}}-D_{\text{KL}}^{\text{cont}}+\lambda L_{\text{HSIC}}\right|+\\&\beta_{\text{disc}}\left|C_{\text{disc}}-D_{\text{KL}}^{\text{disc}}\right|\end{aligned} \tag{5}$$

式中：β_{cont} 和 β_{disc} 分别为连续变量和离散变量的容量权重；C_{cont} 和 C_{disc} 分别为连续通道和离散通道的容量；$D_{\text{KL}}^{\text{cont}}$ 和 $D_{\text{KL}}^{\text{disc}}$ 分别为连续变量和离散变量的 KL 散度。JointVAE 模型通过线性调度策略改变容量权重，实现动态平衡 KL 散度，以此来调整训练不同阶段时的约束强度，并在训练阶段对连续变量组间施加权重为 λ 的 HSIC 正则项损失，来强制连续变量组间参数的独立性，从而提升解耦表征效果。

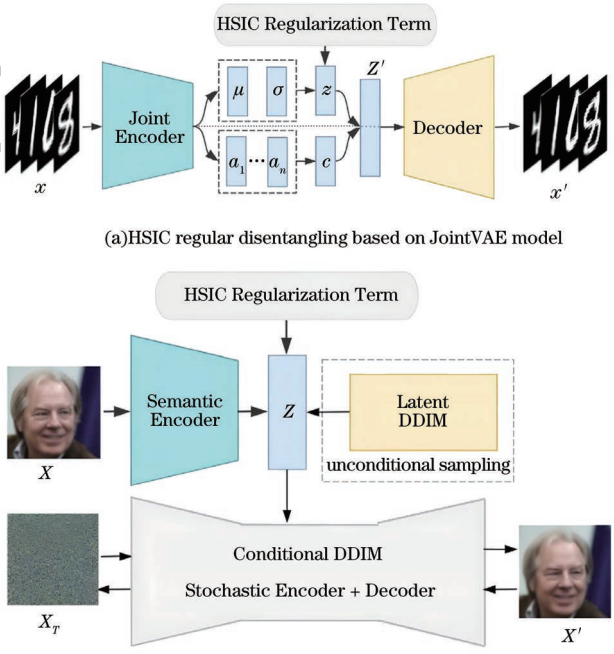


图 2 基于不同模型的解耦改进

Fig. 2 Disentangling improvement based on different models

2.1.2 DM 类模型中的解耦生成

在 DM 类模型中，本研究方法同样将 HSIC 构建为潜在空间独立性正则项，并将其系统性地融入 DM 模型的优化目标中，通过反向过程学习渐进解耦表征。

以基线模型 DiffAE[25] 为例，它通过条件扩散过程与自动编码器的联合设计实现高质量生成与语

义控制，其核心思想是利用扩散过程作为解码器，并通过编码器提取潜在变量以指导生成。在 DiffAE 模型前向扩散过程中，条件扩散模型充当随机编码器，从输入图像 X_0 中捕获图像细粒度细节，弥补由 X_0 编码后语义表示 z 中无法得到的内容。X_0 通过在正向扩散过程中逐步添加噪声，生成噪声序列 $\{X_1, X_2, \cdots, X_T\}$，具体的前向加噪过程为：

$$X_{t+1}=\sqrt{\alpha_{t+1}}f_\theta(X_t,t,z)+\sqrt{1-\alpha_{t+1}}\varepsilon_\theta(X_t,t,z) \tag{6}$$

式中：t 指第 t 次加噪过程；f_θ 和 ε_θ 分别为条件去噪扩散隐式模型（DDIM）的去噪网络和噪声预测模块。在 DiffAE 模型后向扩散过程中，条件 DDIM 作为解码器，对与潜在变量的推理分布对齐的生成过程进行建模，从而确保生成高质量的图像。从噪声 X_T 出发，利用条件 DDIM 逐步去噪生成图像 X_0，条件为语义编码 z，具体的后向去噪过程为：

$$p_\theta(X_{0:T}|z)=p(X_T)\prod_{t=1}^{T}p_\theta(X_{t-1}|X_t,z) \tag{7}$$

通过反向扩散链，模型生成与 z 语义一致的高质量图像。该模型的损失是衡量噪声预测误差的基础，扩散损失为：

$$L_{\text{simple}}=\sum_{t=1}^{T}E_{X_0,\varepsilon_t}[\|\varepsilon_\theta(X_t,t,z)-\varepsilon_t\|_2^2] \tag{8}$$

式中：ε_t 为真实噪声；ε_θ 为模型预测的噪声；E 表示期望；$\|\cdot\|_2^2$ 是 L2 范数的平方，用于衡量预测噪声与真实噪声的差距。

在该基线模型上施加 HSIC 正则的 HSIC-DiffAE 模型，如图 2(b)所示，需要在现有的 DiffAE 模型中对连续潜在变量 z 施加 HSIC，推动各维度间独立，HSIC 正则项值的计算如式(3)所示，总损失如式(9)所示：

$$L_{\text{DiffAE}}=L_{\text{simple}}+\lambda L_{\text{HSIC}} \tag{9}$$

式中：λ 为 HSIC 正则项的权重系数，通过实验调优确定。已知原扩散损失 L_{simple} 通过噪声预测误差约束生成质量，而 HSIC 正则项 L_{HSIC} 则通过隐式约束潜在变量独立性影响生成控制。仅依赖潜在变量的 HSIC 计算不影响扩散核函数的似然计算，即去噪计算如式(7)所示，因此正则项的引入不会破坏扩散模型的马尔可夫链稳态性质。

上述设计表明，基于非参数统计方法 HSIC 建立的潜在表示独立性正则机制，在 VAE 类模型和扩散模型中均可实现，验证了所提方法的理论通用性。

2.2 评估方法

2.2.1 解耦评估方法

为了评估解耦表征效果，现有研究主要依赖两类指标：监督式解耦度量与无监督解耦度量[4]。传统方法多采用监督式度量，需依赖真实语义因子的标注信息[21]，如 DCI[23]、FactorVAE score[22] 和 TAD 指标[24]。然而，此类方法在无标注场景下适用性受限。

本文提出的 HSIC score 属于无监督度量，无需任何先验标签即可基于 HSIC 准则直接量化潜在变量的独立性，并且避免生成干预或预测器训练，计算高效。对连续潜在变量，通过计算各维度与全局统计量间 HSIC 值，评估其独立性。具体来说，基于 n 维潜在变量 z，通过各维度间 HSIC 值的全局统计量均值 z_{mean}，计算 HSIC score(s_{HSIC})：

$$s_{\text{HSIC}}=\frac{1}{n}\sum_{i=1}^{n}\text{HSIC}(z_i,z_{\text{mean}}) \tag{10}$$

计算得到的评估值 HSIC score，该指标越小表明潜在表示维度间的统计独立性越强，解耦效果越好。

2.2.2 潜在空间插值平滑性评估

在潜在变量维度较高或存在复杂非线性依赖时，通过插值生成实验评估解耦效果，常会面临两类典型问题：一是潜在空间插值生成的过渡图像存在不连贯现象，如 DisDiff 指出，若 DM 的反向去噪过程未充分约束潜在变量的独立性，插值生成的图像可能因噪声残留导致局部伪影（如人脸姿态突变或纹理不连续）[8]；二是潜在空间插值生成的图像存在语义不连续现象，且此类问题难以通过视觉量化，如 β-VAE 实验结果表明，若潜在变量间存在强耦合性，插值轨迹会因非线性映射产生不可预测的语义跳跃，仅依赖视觉观察难以准确捕捉此类突变现象[11]。为此，需引入量化指标客观评估潜在空间插值轨迹的平滑性，并结合数值分析与可视化验证，以全面衡量其连续性与解耦性。当前主流的平滑性评估方法为插值标准差（ISTD），该指标由平滑扩散构成，通过统计插值路径的局部线性度衡量潜在空间的连续性[26]。ISTD 值越低，表明插值路径的局部线性度越高，潜在空间越平滑。然而，ISTD 的计算需依赖文本提示生成插值样本，因此在无文本条件的生成任务中，需对其改进以构建非文本依赖的类 ISTD 指标。

在类 ISTD 指标（ISTD2，E_{ISTD2}）计算流程中，首先需要从测试集随机抽取 N 对图像 X_a 和 X_b，分别编码为潜在向量 $\boldsymbol{z}_a$ 和 $\boldsymbol{z}_b$，接着在每对潜在向量间进行线性插值或球面线性插值，生成 k 个中间潜在向量 $\{\boldsymbol{z}_1,\boldsymbol{z}_2,\cdots,\boldsymbol{z}_k\}$，然后将 $k+2$ 个潜在向量解码为重建图像 $\{I_1,I_2,\cdots,I_{k+2}\}$，并计算相邻重建图像

计算机工程 www.ecice06.com

的像素级 L2 距离标准差，最后对图像对的所有插值路径取均值，得到最终评分：

$$E_{\mathrm{ISTD2}}=\frac{1}{N}\sum_{n=1}^{N}\left(\sqrt{\frac{1}{k+1}\sum_{i=1}^{k+1}\left\|I_i^n-I_{i+1}^n\right\|_2^2}\right) \quad (11)$$

该指标通过量化插值路径的平滑性，为潜在空间的连续性与解耦性提供客观评估依据，指导模型优化。

3 实验

3.1 实验设置

本研究实验选用 MNIST[27] 和 CelebA[28] 两个经典数据集，分别验证解耦方法在结构化简单数据与复杂多属性数据下的性能。MNIST 数据集包含 60 000 张训练图像及 10 000 张测试图像，均为 28×28 像素灰度手写数字，用于验证数字形态的独立表征能力。CelebA 数据集包含 202 599 张 RGB 人脸图像，标注 40 种细粒度属性，通过统一中心裁剪至分辨率为 128×128 像素来消除背景干扰，重点评估复杂场景下的细粒度解耦效果。本实验在 PyTorch 框架下实现，硬件环境为单张 NVIDIA RTX 4090 GPU。基线模型涵盖 VAE 和 DM 两类架构，具体包括 VAE、JointVAE 以及 DiffAE 模型。VAE 类模型采用 2 层卷积网络，根据任务潜在维度设为 2 或 10，训练 Epoch 为 100，批量大小为 128，学习率是 0.001。DiffAE 使用 U-Net 主干网络，扩散步长为 1 000，单卡训练批量大小为 32，总训练样本数为 72×10^6（折合约 355 个 Epoch），学习率是 0.000 1。

本研究对基线模型分别施加了 HSIC 正则项，与原模型对比解耦效果，并将改进的 DiffAE 与其它先进模型进行定量对比分析。HSIC 正则项权重系数 λ 经参数调优确定为 100。为了评估解耦效果，本文对生成模型重建任务和潜在空间插值生成进行了实验，并使用了生成质量代表性评估指标 FID，解耦度量指标 DCI、FactorVAE score 和 TAD、潜在空间平滑性指标 ISTD2，以及所提的解耦指标 HSIC score。所有实验需独立重复 3 次以消除随机性。

3.2 生成质量与解耦效果

3.2.1 生成质量

在 MNIST 数据集上，本文对原始 VAE 与 HSIC 正则化的 HSIC-VAE 模型进行重建能力对比实验。定性结果如图 3 所示，第 1 行是原始输入图像，第 2 行是 VAE 原模型重建结果，第 3 行是 HSIC-VAE 模型重建结果。从图 3 可以看出，HSIC-VAE 模型的重建图像与原模型在视觉上无明显差别，表明 HSIC 正则化未显著影响模型局部细节。



图 3 VAE 和 HSIC-VAE 模型的原图重建结果

Fig.3 Original image reconstruction results of VAE and HSIC-VAE models

实验使用 FID 来评估生成质量，FID 值越低，模型生成质量越高。定量结果如表 1 所示，加粗表示最优数据（下同）。从表 1 可以看出，通过计算两模型的 FID 评分，发现 HSIC-VAE 的 FID 取值较 VAE 升高了 7 百分点，表明 HSIC 正则化使模型生成的图像质量变差。HSIC 正则项的引入使得模型质量评估 FID 变差和其他解耦指标变好，这本质上反映了解耦目标与生成质量间的内在权衡，独立性正则化约束虽然优化了解耦效果，但可能破坏数据中固有属性相关性，导致生成图像细节保真度下降。同时，新增的 HSIC 正则项改变了 VAE 的优化目标，原本集中于重构精度和分布对齐的模型容量被部分分散到解耦约束优化，导致生成图像的整体分布偏离真实数据。FID 差值虽然远超出误差范围，但因其结果值较大导致整体统计不显著，不会显著影响生成质量。这初步证明了 HSIC 正则化方法可以控制生成质量的损失。

表 1 VAE 和 HSIC-VAE 模型的对比评估度量

Table 1 Comparative evaluation metrics of VAE and HSIC-VAE models

Model	FID ↓	HSIC score ↓	Factor VAE score ↑	DCI ↑	ISTD2 ↓
VAE	**134.4**	200.4	0.917	0.177	0.259
HSIC-VAE	141.4	**115.3**	**0.921**	**0.569**	**0.235**

在 CelebA 数据集上，本文将原始 DiffAE 与 HSIC 正则化的 HSIC-DiffAE 模型进行重建能力对比实验。定性结果如图 4 所示，HSIC-DiffAE 模型的重建图像与原模型在视觉上几乎一致，同样表明 HSIC 正则化未显著影响模型局部细节。

在定量结果评估方面，不仅将该 HSIC-DiffAE 模型与 DiffAE 模型进行重建能力对比，还增加了其他多个基线模型参与评估，结果如表 2 所示。HSIC-DiffAE 模型的 FID 数值略高于 DiffAE 原模型，两模型的 FID 差异为 0.7 百分点，牺牲数值未达到统计显著性，表明HSIC正则化的引入可能轻

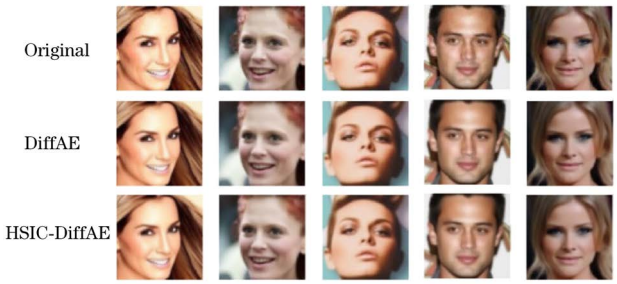


图 4 DiffAE 和 HSIC-DiffAE 模型的原图重建结果

Fig. 4 Original image reconstruction results of DiffAE and HSIC-DiffAE models

微牺牲部分生成质量，但两者 FID 值显著低于其他基线模型，证明 HSIC 正则化仍然保持着原模型的高生成质量，说明 HSIC 对生成质量的影响较小，包括 CelebA 这样的复杂数据。

表 2 在 CelebA 数据集上各模型的评估度量

Table 2 Evaluation metrics for various models on the CelebA dataset

Model	FID↓	TAD↑
VAE	94.3±2.9	0
β-VAE	99.8±2.4	0.088±0.051
InfoVAE	77.8±1.6	0
DiffAE	**22.7±2.1**	0.155±0.010
HSIC-DiffAE	23.4±2.2	**0.244±0.008**

模型生成图像的质量需要从逼真性及多样性方面进行评估。由于图像重建实验仅能验证图像保真度，为量化多样性表现，因此本研究在 CelebA 数据集上开展了 HSIC-DiffAE 模型的随机采样实验。具体可视化采样示例如图 5 所示。HSIC-DiffAE 模型的采样图像与原模型类似，在保留细节的同时，涵盖肤色、发型、姿态等多样化特征，表明 HSIC 正则化的模型能保证生成图像的逼真性，且仍然维持多样性。

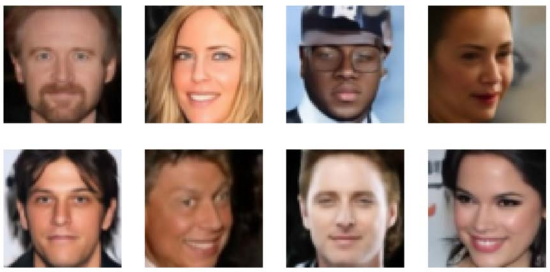


图 5 HSIC-DiffAE 模型的图像采样结果

Fig. 5 Image sampling results of HSIC-DiffAE model

总体而言，在图像生成质量方面，HSIC 正则化对模型生成结果的影响具有可控性。与原模型相比，经 HSIC 正则化的模型仍能保持逼真性与多样性。

3.2.2 模型解耦效果

为验证 HSIC 正则化方法对解耦效果的影响，在 MNIST 数据集上，本文分别对潜在维度为 10 的 VAE 及 HSIC-VAE 模型进行潜在空间插值实验，对 VAE 和 HSIC-VAE 模型中每个潜在变量维度分别进行独立插值，观察每一维度改变时图像属性的变化，定性结果如图 6 所示，每一行代表图像编码潜在变量的单一维度插值变化结果图。结果经对比发现，图 6(b)所示为 HSIC-VAE 模型的属性解耦更明显，呈现出分离的数字倾斜属性和数字类型属性，而 VAE 结果存在属性混淆现象，表明了 HSIC 正则化具有较好的解耦效果，能有效分离语义因子。

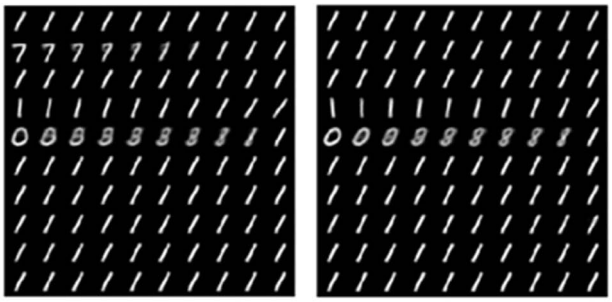


图 6 VAE 和 HSIC-VAE 模型的解耦效果

Fig. 6 Disentangling effect of VAE and HSIC-VAE models

实验使用 HSIC score、FactorVAE score 和 DCI 值来评估解耦效果，HSIC score 值越低，FactorVAE score 和 DCI 值越高，模型的解耦效果就越好。定量结果如表 1 所示，两模型的解耦度量值符合图像定性结果，HSIC-VAE 模型具有更优的解耦度量值，表明 HSIC 正则化促进解耦表征。具体来讲，HSIC-VAE 模型具有更低的 HSIC score，表明模型潜在变量间更独立，它具有更高的 FactorVAE score 和 DCI 值，则展现出了潜在变量与语义属性间的更强关系，可解释性强。

在 CelebA 数据集上，针对 DiffAE 和 HSIC-DiffAE 模型开展了多维度潜在向量插值实验。具体操作从模型编码得到的 512 维潜在向量中选取 10 维，以 10 步长进行线性插值。实验结果如图 7 所示。从图 7 可以看出，HSIC-DiffAE 模型能够通过这 10 维控制人眼状态，而原模型则无法实现这一控制。进一步分析表明，HSIC 正则化促使模型向量尽可能保持独立，使选择的 10 维向量能够有效映射到人眼特征，而原模型未能建立这种映射关系。实验结果表明，HSIC 正则化具备一定的解耦效果，对复杂数据具有一定的可控性。

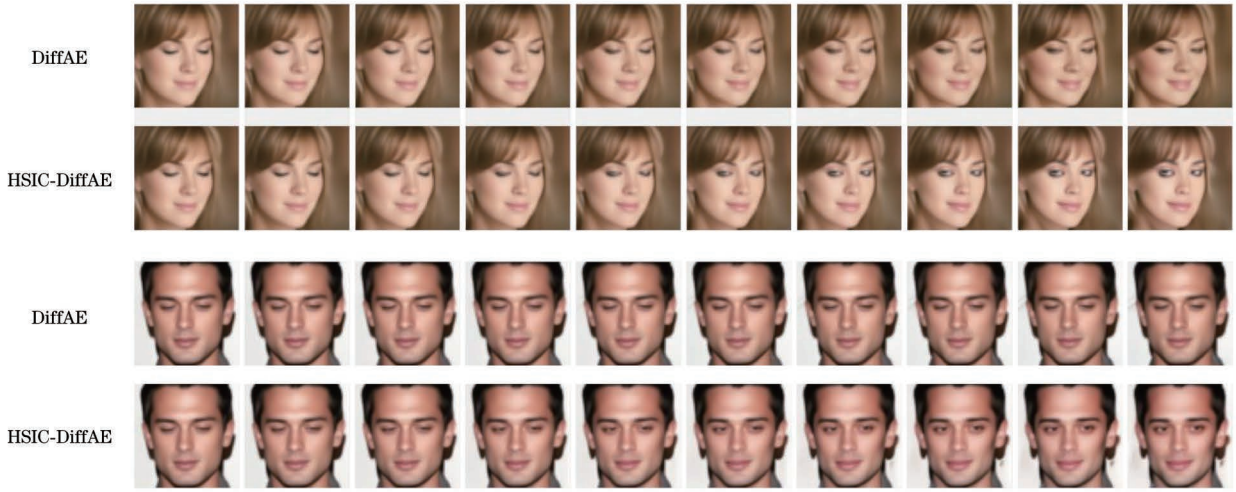


图 7 DiffAE 和 HSIC-DiffAE 模型的解耦效果

Fig. 7 Disentangling effect of DiffAE and HSIC-DiffAE models

在定量结果评估方面，不仅将该 HSIC-DiffAE 模型与 DiffAE 模型进行解耦能力对比，还纳入了其他多个基线模型参与评估。因逻辑适配，CelebA 数据集常采用 TAD 指标进行解耦度量，TAD 值越高，解耦效果越好。通过 TAD 度量值来定量评估解耦效果，实验结果如表 2 所示，HSIC-DiffAE 在测量模型中具有最高的 TAD 值，表明其具有最好的解耦效果。从表 2 中还可以看出，VAE 和 InfoVAE 两个模型的 TAD 值为 0，原因是它们的潜在变量未能解耦，导致每个属性的最高和次高 AUROC 分数差异不足。

总的来说，不同实验 FID 的轻微上升伴随着解耦指标的显著提升，证明了所提方法可以在质量损失可控范围下实现高效解耦。

3.3 潜在空间平滑性插值实验

本研究在 3 组对比模型实验中进行潜在空间平滑性插值实验，即在 MNIST 数据集上的 VAE 和 HSIC-VAE、JointVAE 和 HSIC-JointVAE 以及在 CelebA 数据集上的 DiffAE 和 HSIC-DiffAE 三组模型进行实验。

3.3.1 平滑性度量

在潜在空间平滑性评估方面，本文采用前面提出的 ISTD2 度量。ISTD2 值越低，表明插值路径的局部线性度越高，潜在空间越平滑。3 组插值实验的定量结果，如表 1、表 3 和表 4 所示，具有 HSIC 正则项的模型都具有更低的 ISTD2 值，表明 HSIC 正则化通过减少隐变量冗余关联，增强了潜在空间的平滑性。基于 VAE 架构的对比模型间具有更大的差值，表明平滑增强效果更强。基于 DM 架构的对比模型，相较于 VAE 架构，具有更高的 ISTD2 值，表明基于 VAE 架构的模型插值生成结果更加平滑。

表 3 JointVAE 和 HSIC-JointVAE 模型的对比评估度量

Table 3 Comparative evaluation metrics of JointVAE and HSIC-JointVAE models

Model	FID↓	HSIC score↓	Factor VAE score↑	DCI↑	ISTD2↓
JointVAE	**119.2**	46.00	0.945	0.169	0.263
HSIC-JointVAE	128.1	**35.00**	**0.948**	**0.599**	**0.231**

表 4 DiffAE 和 HSIC-DiffAE 模型的对比评估度量

Table 4 Comparative evaluation metrics of DiffAE and HSIC-DiffAE models

Model	FID↓	HSIC score↓	Factor VAE score↑	TAD↑	ISTD2↓
DiffAE	**22.7**	0.23	0.485	0.155	0.377
HSIC-DiffAE	23.4	**0.20**	**0.492**	**0.244**	**0.373**

3.3.2 低维潜在空间插值

在低维潜在空间插值定性实验方面，先是对 VAE 和 HSIC-VAE 模型进行全维度及单维度插值实验。实验结果如图 8 所示，图 8(a)为 VAE 插值结果，图 8(b)为 HSIC-VAE 插值结果，第 1 行是潜在变量全维度同时插值变化结果，第 2 行是潜在变量第 1 个维度插值变化结果，第 3 行是潜在变量第 2 个维度插值变化结果。例如从左上角图数字 1 向右下角图数字 7 的连续 2 维潜在变量插值。HSIC-VAE 的单维插值生成图像分别在数字类型、笔画粗细的属性变化上表现出更强的独立性，例如图 8(b)中第 2 行的单一维度插值实验，能够精准分离并改变单一属性笔画粗细，而其他属性(如倾斜角度、数字类别等)保持稳定，表明 HSIC 正则化有效提升了潜在变量的解耦能力。HSIC-VAE 的整体插值生成图像则表现出变化的平

滑性，如第 1 行的实验结果中，HSIC-VAE 没有出现过多的其他数字类别伪影。

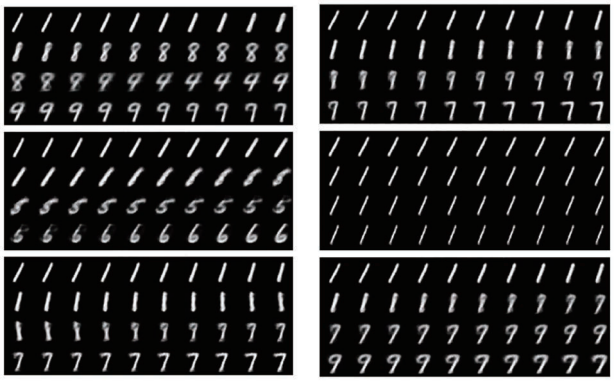


图 8 VAE 和 HSIC-VAE 模型低维潜在空间插值结果

Fig. 8 Low dimensional latent space interpolation results of VAE and HSIC-VAE models

接着，在 JointVAE 和 HSIC-JointVAE 模型的 10 维潜在变量插值生成实验中，如图 9 所示，两对比模型分别对 0～9 这 10 个数字类别的一组数字图像的单维潜在变量进行均匀插值，生成新的 10×10 张图片。从图 9 可以看出，HSIC-JointVAE 模型的插值生成效果显著优于 JointVAE，不仅将笔画粗细属性独立表征，而且不影响数字类别，但 JointVAE 则存在属性耦合现象，即粗细变化伴随数字类型偏移。

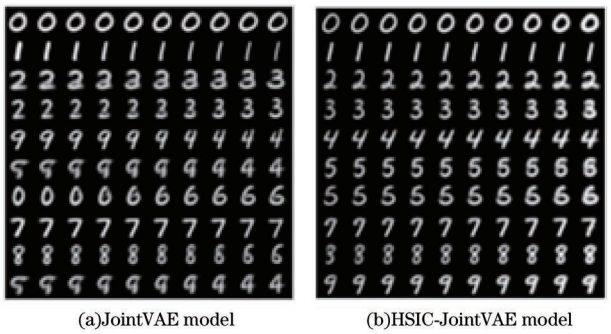


图 9 JointVAE 和 HSIC-JointVAE 模型低维潜在空间插值结果

Fig. 9 Low dimensional latent space interpolation results of JointVAE and HSIC-JointVAE models

3.3.3 高维潜在空间插值

在高维潜在空间插值定性实验方面，根据 CelebA 数据集上的数据进行 512 维潜在变量插值。实验结果如图 10 所示，在从左图人脸插值生成右图人脸的实验中发现，HSIC-DiffAE 模型展现出更平滑的细节过渡和更少的伪影生成，如第 2 行头发的稀疏渐变和第 4 行嘴角的自然渐变，表明 HSIC 正则化提升了潜在空间的连续性。



图 10 DiffAE 和 HSIC-DiffAE 模型高维潜在空间插值结果

Fig. 10 High dimensional latent space interpolation results of DiffAE and HSIC-DiffAE models

综上，HSIC 正则化方法通过约束潜在变量独立性，不仅显著提升了模型的解耦性能，还增强了潜在空间的平滑性。

3.4 对比消融实验

为了进一步验证 HSIC 正则化这一解耦方法的有效性，本文开展多个模型下的对比消融研究实验。实验在 VAE、JointVAE 和 DiffAE 3 个基线模型上进行 HSIC 正则项的消融研究，具体实验结果如表 1、表 3 和表 4 所示，每张表上包含有无 HSIC 正则项的模型对比结果。在生成质量评估方面，每个对比结果中都是采用了 HSIC 正则方法的模型具有略高的 FID，表明正则化解耦优化可能牺牲部分生成细节，但是两者数值接近，说明 HSIC 正则化方法对生成质量的影响较小。在解耦表征评估方面，含 HSIC 正则项的模型都具有较高的 FactorVAE score、DCI 或 TAD 值，验证了 HSIC 正则化方法对属性解耦表征的有效性和通用性，更低的 ISTD2 值则表明含 HSIC 正则项的模型插值生成更加平滑。同时测量

评估了模型的 HSIC score 指标，结果发现正则优化后的模型具有更低的 HSIC score 值，符合评估原理，进一步验证了 HSIC score 指标对解耦能力评估的有效性。

综上，通过 MNIST 与 CelebA 数据集上的系统实验，HSIC 正则化方法在 VAE、JointVAE、DiffAE 3 个模型中，均展现出稳定的解耦性能提升，明显的潜在空间平滑性增强，且对生成质量的负面影响有限，这证实了其具有跨模型架构的通用性，并能平衡解耦优化与生成保真度的任务需求。

4 结束语

针对生成模型在没有标注数据的图像生成场景下特征独立性表征的明显缺陷，本研究提出了一种基于 HSIC 正则化的通用解耦生成方法。通过将 HSIC 度量转化为潜在空间的正则化项，在强制潜在变量独立性的同时平衡重构损失，确保无监督训练场景下的稳定性。理论分析表明，该方法可以有效建模特征间的非线性及高阶统计依赖，为解耦学习提供了新的优化视角。实验结果表明，该方法在 VAE 和 DM 两类主流生成模型架构中均表现出优异的解耦性能，显著提升了潜在表示的独立性，且优化了插值轨迹的平滑性。本研究的核心创新点在于，将 HSIC 这一非参数统计方法转化为正则项，系统地融入生成模型的优化目标中，为生成模型解耦提供统一框架。此外，本文所提的评估指标 HSIC score，弥补无监督解耦量化评估的空白。在不同模型架构中的实验证明了该方法的通用性和鲁棒性，为生成模型在复杂场景中的应用奠定了理论基础。后续将探索 HSIC 在多模态生成模型中的应用，以解决跨模态特征对齐的难题。此外，将该方法应用于实际场景，如图像修复或水下图像增强等，有助于探索其在实际问题中的实用价值。

参考文献

[1] 蔡江海，黄成泉，王顺霞，等. 基于解耦表征学习的生成式视觉图像理解[J]. 计算机辅助设计与图形学学报，2025，37(12)：2202-2212.
CAI J H，HUANG C Q，WANG S X，et al. Generative visual image understanding based on disentangled representation learning [J]. Journal of Computer Aided Design and Graphics，2025，37(12)：2202-2212. (in Chinese)

[2] 文载道，王佳蕊，王小旭，等. 解耦表征学习综述[J]. 自动化学报，2022，48(2)：351-374.
WEN Z D，WANG J R，WANG X X，et al. A review of disentangling representation learning [J]. Journal of Automation，2022，48 (2)：351-374. (in Chinese)

[3] BENGIO Y，COURVILLE A，VINCENT P. Representation learning：a review and new perspectives [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2013，35(8)：1798-1828.

[4] WANG X，CHEN H，TAO S A，et al. Disentangled representation learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2024，46 (12)：9677-9696.

[5] HIGGINS I，MATTHEY L，PAL A，et al. β-VAE：learning basic visual concepts with a constrained variational framework [EB/OL]. [2025-02-20]. https：//openreview.net/pdf?id=Sy2fzU9gl.

[6] CHEN R T Q，LI X C，GROSSE R，et al. Isolating sources of disentanglement in variational autoencoders [EB/OL]. [2025-02-20]. https://arxiv.org/abs/1802.04942.

[7] CHEN X，DUAN Y，HOUTHOOFT R，et al. InfoGAN：interpretable representation learning by information maximizing generative adversarial nets [EB/OL]. [2025-02-20]. https://arxiv.org/abs/1606.03657.

[8] YANG T，WANG Y W，LV Y，et al. DisDiff：unsupervised disentanglement of diffusion probabilistic models [EB/OL]. [2025-02-20]. https://arxiv.org/abs/2301.13721.

[9] 曾文献，张曼钰，孙磊. 生成对抗式双重解耦的分阶段阴影去除算法[J]. 计算机工程与应用，2025，61(20)：260-269.
ZENG W X，ZHANG M Y，SUN L. Generative adversarial double disentangling staged shadow removal algorithm [J]. Computer Engineering and Applications，2025，61(20)：260-269. (in Chinese)

[10] 刘彦呈，董张伟，朱鹏莅，等. 基于特征解耦的无监督水下图像增强[J]. 电子与信息学报，2022，44(10)：3389-3398.
LIU Y C，DONG Z W，ZHU P L，et al. Unsupervised underwater image enhancement based on feature disentangling [J]. Chinese Journal of Electronics and Information Technology，2022，44 (10)：3389-3398. (in Chinese)

[11] BURGESS C P，HIGGINS I，PAL A，et al. Understanding disentangling in β-VAE [EB/OL]. [2025-02-20]. https：//arxiv.org/pdf/1804.03599.

[12] VAN DEN OORD A，VINYALS O，KAVUKCUOLU K. Neural discrete representation learning [EB/OL]. [2025-02-20]. https://arxiv.org/pdf/1711.00937.

[13] KARRAS T，LAINE S，AILA T. A style-based generator architecture for generative adversarial networks [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C.，USA：IEEE Press，2019：4401-4410.

[14] CAO H Q，TAN C，GAO Z Y，et al. A survey on generative diffusion models [J]. IEEE Transactions on Knowledge and Data Engineering，2024，36(7)：2814-2830.

[15] GRETTON A，BOUSQUET O，SMOLA A，et al. Measuring statistical dependence with Hilbert-Schmidt norms [C] // Proceedings of International Conference on Algorithmic Learning Theory. Berlin，Germany：Springer，2005：63-77.

[16] DUPONT E. Learning disentangled joint continuous and discrete representations [EB/OL]. [2025-02-20]. https：//arxiv.org/pdf/1804.00104.

[17] KIM Y，KIM D，LEE H M，et al. Unsupervised controllable generation with score-based diffusion models：disentangled latent code guidance [C] // Proceedings of Conference on Neural Information Processing Systems. New York，USA：Neural Information Processing Systems Foundation，2022：25-31.

[18] MA W D K，LEWIS J P，KLEIJN W B. The HSIC bottleneck：deep learning without back-propagation [C] // Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto，USA：AAAI Press，2020：5085-5092.

[19] LIU X，THERMOS S，SANCHEZ P，et al. HSIC-InfoGAN：learning unsupervised disentangled representations

by maximising approximated mutual information [C] // Proceedings of MICCAI Workshop on Medical Applications with Disentanglements. Berlin, Germany: Springer, 2022: 15-21.
[20] WANG Z F, ZHAN Z, GONG Y F, et al. DualHSIC: HSIC-bottleneck and alignment for continual learning [EB/OL]. [2025-02-20]. https://arxiv.org/abs/2305.00380.
[21] CARBONNEAU M A, ZAIDI J, BOILARD J, et al. Measuring disentanglement: a review of metrics [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 35(7): 8747-8761.
[22] KIM H, MNIH A. Disentangling by factorising [EB/OL]. [2025-02-20]. https://arxiv.org/abs/1802.05983.
[23] EASTWOOD C, WILLIAMS C K I. A framework for the quantitative evaluation of disentangled representations [C] // Proceedings of the 6th International Conference on Learning Representations. New York, USA: [s.n], 2018: 65-69.
[24] YEATS E, LIU F, WOMBLE D, et al. NashAE: disentangling representations through adversarial covariance minimization [C] // Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2022: 36-51.
[25] PREECHAKUL K, CHATTHEE N, WIZADWONGSA S, et al. Diffusion autoencoders: toward a meaningful and decodable representation [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D.C., USA: IEEE Press, 2022: 10619-10629.
[26] GUO J Y, XU X Q, PU Y F, et al. Smooth diffusion: crafting smooth latent spaces in diffusion models [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D.C., USA: IEEE Press, 2024: 7548-7558.
[27] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
[28] LIU Z W, LUO P, WANG X G, et al. Deep learning face attributes in the wild [C] // Proceedings of the IEEE International Conference on Computer Vision. Washington D.C., USA: IEEE Press, 2015: 3730-3738.

文字编辑 薛晋栋
栏目编辑 宋 圆

计算机工程 www.ecice06.com