基于多特征融合与渐进式学习的SM4密码算法差分区分器

王慧娇，张哲
（桂林电子科技大学计算机与信息安全学院，广西 桂林 541004）

摘　要：随着深度学习在密码分析中的深入应用，现有区分器仅依赖单–特征已无法有效区分具有强非线性扩散且高轮次的SM4密码算法。为突破SM4神经差分区分器在特征提取能力和高轮次上区分性能的瓶颈，提出一种基于多特征融合与渐进式学习的神经区分器模型。设计高阶交叉差分密文对预处理方法增强非线性特征密度，结合Inception架构多核卷积层及SE通道注意力机制强化特征捕获，并采用渐进式学习策略阶段化训练区分器模型。实验结果表明，所提模型在5~9轮区分的准确率分别提升至75.12%、64.90%、60.62%、56.13%和53.88%；首次实现10轮SM4神经差分区分器（测试集准确率达52.41%），突破当前已知的9轮区分轮数瓶颈；对模型的特异度与敏感度进行测试分析，揭示了模型更擅长“排除”非密文的区分特点，验证了多特征融合及渐进式学习在SM4密码算法安全性评估中的有效性和可行性。

关键词：分组密码；差分分析；SM4密码算法；神经差分区分器；多特征融合
中图分类号：TN918
文献标志码：A
DOI：10.11959/j.issn.1000–436x.2025250

Multi-feature fusion and progressive learning-based differential distinguisher for SM4 cipher algorithm

WANG Huijiao, ZHANG Zhe
School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China

Abstract: As deep learning advanced in cryptanalysis, existing distinguishers relying on single features are ineffective for the highly nonlinear and strongly diffusive SM4 cipher at high encryption rounds. To address the limitations in feature extraction capability and high-round distinction performance of SM4 neural differential distinguisher, a neural distinguisher model based on multi-feature fusion and progressive learning was proposed. A high-order cross-differential ciphertext pair preprocessing method was designed to enhance nonlinear differential characteristics. An Inception-based multi-kernel convolution structure integrated with a squeeze-and-excitation channel attention mechanism improves feature capture. A progressive learning strategy was adopted to train the distinguisher model in stages. Experimental results show that the proposed model respectively achieves classification accuracy improvements of 75.12%, 64.90%, 60.62%, 56.13%, and 53.88% on rounds 5 through 9.For the first time, a 10-round SM4 neural differential distinguisher is realized (with a test set accuracy of 52.41%), breaking through the current known bottleneck of the number of rounds for 9-round distinction. The test and analysis of the specificity and sensitivity of the model reveals the distinction feature that the model is better at “excluding” non-ciphertext, validating the effectiveness and feasibility of multi-feature fusion and progressive learning in assessing the security of the SM4 cipher.

Keywords: block cipher, differential analysis, SM4 cipher algorithm, neural differential distinguisher, multi-feature fusion

收稿日期：2025–06–26；**修回日期：**2025–12–01
基金项目：国家自然科学基金资助项目（No.62402132）；广西科技重大专项基金资助项目（No.桂科AA.22068072）；桂林电子科技大学研究生教育创新计划基金资助项目（No.2024YCXS062）
Foundation Items: The National Natural Science Foundation of China (No.62402132),The Guangxi Science and Technology Major Program (No.Guike AA.22068072), The Innovation Project of GUET Graduate Education (No.2024YCXS062)

0 引言

分组密码作为主流的对称密钥密码原语之一，凭借其高效性、可证明安全性及标准化的实现特性，在数据的加密传输与加密存储领域占据核心地位。该技术通过将明文划分为固定长度的分组单元，并基于密钥控制的置换–代换网络实现加密变换，既能保障数据传输的机密性与完整性，又能满足实时通信场景下的吞吐量需求。SM4密码算法作为我国自主设计并标准化的商用分组密码之一，已在政务、金融及通信等领域中承担核心数据加密职能，其安全性与应用可靠性具有重要研究价值。

当前广泛使用的分组密码安全性分析方法主要包括差分分析[1]、积分分析[2]、线性分析[3]等。其中，差分分析由Biham等[4]提出，是分组密码攻击中最有力的工具之一。作为差分分析的核心工具，差分区分器通过刻画明文差分与密文差分之间的统计关联，为密钥恢复攻击提供区分基础。然而，随着计算能力的持续提升以及密码分析理论与方法的不断演进，分组密码体系在复杂应用环境下面临着愈发多样化的攻击模型与安全挑战，发展新的密码分析技术是当前密码学的重点研究方向之一。

近年来，人工智能技术取得突破性进展，以深度神经网络为代表的深度学习（DL, deep learning）方法已在多个领域展现出出色的表征学习能力，包括自然语言处理[5]、医学影像辅助诊断[6]、图像处理[7]、自动驾驶[8]、生物信息[9]、对抗样本防御[10]以及生物特征隐私保护[11]等应用场景，DL方法均实现了传统方法难以企及的性能提升。这种技术革命同样渗透至密码学领域——早在机器学习理论发展初期，Rivest等[12]便从计算复杂性视角揭示了密码系统与机器学习模型的内在关联性，其研究指出：密码算法设计的扩散机制与神经网络的特征提取过程具有数学同构性，这为构建基于DL的密码分析框架奠定了理论基础。Gohr[13]首次采用深度学习对黑箱密码进行分析，将密文与随机置换的判别问题转化为分类任务，并通过神经网络训练实现了高效识别。这种基于梯度优化的密钥空间搜索方法，突破了传统密码分析对人工特征工程的依赖，标志着神经密码分析（NC, neural cryptanalysis）研究范式的正式确立。2022年，Baksi[14]使用神经网络设计出GIMLI、ASCON、KNOT和CHASKEY4种密码的神经差分区分器。Gohr等[15]展示了针对代表5种密码结构的6种不同类型的密码进一步优化的神经网络，该网络针对不同密码体系进行了专门设计，同时深入探讨了神经网络预测准确度与差分分布之间的内在关联。2024年，Yadav等[16]提出了一种基于最大似然比的高精度机器学习区分器模型专门应对大分组密码，该模型采用多层感知机（MLP, multi layer perceptron）作为神经网络架构，包含输入层、输出层和2层隐藏层，通过结合低精度的神经差分区分器与增加数据量、使用阈值概率和数据复杂度参数来优化精度。研究表明，在GIFT-128的7轮加密和ASCON的4轮加密下，准确率分别达到了98.8%和99.4%，并且在2^{18}的数据复杂度条件下成功构建出一个针对GIFT-128的8轮神经差分区分器，其准确率高达99.8%。2025年，Yuan等[17]提出基于多差分的相关密钥神经差分区分器框架，通过改进遗传优化算法自动搜索高偏差输入–密钥差分组合，并采用多差分样本构造策略从多条有效差分生成正负样本，丰富模型训练特征。实验证明，该方法在3个SIMECK变体和10个SIMON变体的多轮测试中，显著提升神经差分区分器准确率，将SIMON可区分轮次最高扩展至17轮；同时首次构建7种未研究的SIMON变体区分器，验证了方法的通用性与优越性。

SM4密码算法于2021年6月正式被选为国际标准算法，成为中国无线局域网认证和隐私基础设施国家密码行业标准[18]，被广泛应用于中国的信息安全领域。该算法采用32轮非平衡Feistel结构，其S盒设计遵循严格的双射特性与非线性优化准则，在抗差分/线性攻击方面达到工业级安全标准。随着算法标准化进程的推进，针对SM4密码算法的轮数缩减分析研究逐渐成为密码学界的热点方向。Zhang等[19]构建了5轮迭代差分特征（平均概率为2^{-42}），并基于矩形攻击框架对16轮SM4密码算法实施有效攻击，同时发现12条18轮有效差分路径。Su等[20]通过改进的活跃S盒下界推导方法，确定6、7、12轮SM4密码算法的最小活跃S盒数分别为3、5、12，并成功构造19轮差分特征（约2^{14}条），最终利用2^{118}选择明文实现23轮SM4密钥恢复攻击。潘印雪等[21]提出了针对8 bit S盒的混合

整数线性规划（MILP, mixed integer linear programming）建模方法，构建包含非线性约束的优化模型，从而实现了S盒差分传播特性的精准描述。王敏等[22]在SM4密码算法加密过程的最后4轮中刻意注入故障，缩小迭代轮次，并利用生成的错误密文成功恢复完整的轮密钥信息。文献[23]使用残差神经网络模型建立神经差分区分器，成功训练得到了4~9轮的SM4神经差分区分器，所获得区分器的复杂度和准确率远优于传统差分区分器。毛永霞等[24]使用增加相关分支输入可分性之间的约束条件和对包含非独立可分性传播的矩阵限制非独立传播比特2种基于比特可分性的建模策略，构造了SM4密码算法的13轮积分区分器。

综上所述，深度学习在分组密码安全性研究领域已取得了显著进展，但在面向长分组、长密钥的密码算法时仍面临关键挑战。以SM4密码算法为代表的典型32轮分组密码中，随着加密轮次的不断迭代，初始差分会在S盒与线性变换的双重扩散作用下迅速衰减并呈现高度稀疏化，导致可利用的有效差分信息显著减少。差分特征的急剧弱化直接导致深度学习模型难以捕获高轮次加密后的微弱统计偏差，进而使区分精度在高轮次加密出现明显下降，限制了神经差分区分器在真实安全性评估场景中的适用性。

因此，如何在SM4密码算法的高轮次加密过程中增强区分器模型对隐含差分特征的挖掘与表征能力，成为当前深度学习驱动密码分析研究中亟待突破的核心问题。针对这一挑战，本文提出基于多特征融合与渐进式学习的神经差分区分器模型，用于SM4区分器密文与随机数据的区分任务。本文主要贡献如下。

1) 提出一种增强差分特征密度的高阶交叉差分密文对预处理方法，该方法通过将密文数据转化为高阶交叉差分密文对的输入格式，有效增强了非线性特征的表征密度。

2) 提出一种多特征融合的SM4神经差分区分器模型，在Tensorflow框架下引入GoogLeNet网络[25]中的Inception架构的多核卷积层与SE通道注意力机制实现差分路径的关联分析，提高了神经差分区分器特征提取能力。

3) 在训练SM4神经差分区分器模型时，采用渐进式训练策略突破模型轮次瓶颈，首次构建出有效的10轮SM4神经差分区分器。

1　预备知识

1.1　差分分析及神经差分区分器

差分密码分析是迭代密码的选择明文攻击技术，其核心逻辑为通过构造特定明文差分、捕捉对应密文差分的统计特征，即找到一条高概率差分路径，进而实现密钥信息的高效恢复。该技术是评估分组密码安全性的关键手段，也是本文开展相关密码算法安全性分析的重要理论基础。

定义1　差分和差分对。对于2个l比特流$\boldsymbol{X}\in B^l$和$\boldsymbol{X}^*\in B^l$，它们的异或差分$\Delta\boldsymbol{X}$定义为$\Delta\boldsymbol{X}=\boldsymbol{X}\oplus\boldsymbol{X}^*$。对于固定的差分$\Delta\boldsymbol{X}$，存在恰好$2^l$对$(\boldsymbol{X},\boldsymbol{X}^*)$，使得$\Delta\boldsymbol{X}=\boldsymbol{X}\oplus\boldsymbol{X}^*$。差分对被定义为2个差分的组合，即$(\Delta\boldsymbol{X},\Delta\boldsymbol{Y})$。

定义2　差分路径。r轮差分路径$\boldsymbol{\Omega}$是由明文差分与各轮输出差分构成的差分向量，其定义为

$$(\alpha_0,\alpha_1,\cdots,\alpha_r) \tag{1}$$

其中，α_i是第$i(1\leqslant i\leqslant r)$轮输出$\boldsymbol{X}_i$和$\boldsymbol{X}_i^*$的差分。

定义3　差分概率。在r轮差分路径$\boldsymbol{\Omega}=\alpha_0,\alpha_1,\cdots,\alpha_r$中，差分概率描述该路径中第$i$轮特定差分的出现概率，定义为

$$p_i^{\boldsymbol{\Omega}}=P(\Delta F(\boldsymbol{X})=\alpha_i|\Delta\boldsymbol{X}=\alpha_{i-1}) \tag{2}$$

定义4　差分路径概率。r轮差分路径$\boldsymbol{\Omega}=\alpha_0,\alpha_1,\cdots,\alpha_r$的概率$p^{\boldsymbol{\Omega}}$则表示整个差分向量的整体发生概率，定义为

$$p^{\boldsymbol{\Omega}}=\prod_{i=1}^{r}p_i^{\boldsymbol{\Omega}} \tag{3}$$

神经差分区分器实现真实密文与随机密文的区分与传统差分分析不同，其采用深度学习方法，利用其强特征提取与模式识别能力构造区分器模型，由Gohr于2019年首次提出[13]。神经差分区分器模型的本质为二分类神经网络模型，以真实密文对与随机数据的正确分类比例（准确度）衡量性能，通常以准确度超过0.5视为有效。

1.2　SM4密码算法

SM4密码算法采用128 bit分组与128 bit密钥，基于广义非平衡Feistel网络结构实现加密与解密运算，其加解密流程一致，仅解密时轮密钥需按加密密钥的逆序使用。SM4密码算法加密结构如图1所示。

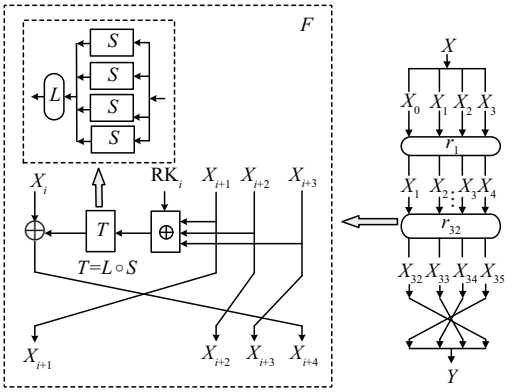

图 1 SM4 密码算法的加密结构

SM4 密码算法的核心组件为合成置换函数 T，其由非线性变换 τ 和线性变换 L 级联构建。其中，非线性变换 τ 由 4 个并行的 8 bit S 盒构成，完成 8 bit 输入字节到 8 bit 输出字节的非线性映射；线性变换 L 则对非线性变换 τ 的输出执行 32 bit 循环移位，实现数据的线性扩散。明文经过 32 轮迭代加密后，需通过反序变换后输出密文。SM4 加密算法的核心安全特性集中于强非线性 S 盒和线性扩散的协同设计，这使差分传播呈高维耦合结构特征。

2 多特征融合的 SM4 神经差分区分器

受 Benamira 等[26]的工作以及深度学习中数据增强技术的启发，本文从数据预处理和提升区分器捕获特征能力 2 个角度出发，进一步优化神经差分区分器的区分性能。设计了一种采用高阶交叉差分密文对预处理方法以增强非线性特征密度的多特征融合神经差分区分器模型。下文将系统阐述区分器模型的解耦理论与构建过程，包括数据预处理、模型结构、数据生成以及模型训练和验证。

2.1 特征解耦与协同表达

在 SM4 密码算法中，由于 S 盒与线性变换共同作用使差分特征传播过程中呈现非线性耦合关系，因此本文在 SM4 神经差分区分器模型中引入基于特征解耦与协同表达理论的多特征融合学习策略。该策略旨在从多尺度、多通道视角充分挖掘差分传播过程中的局部与全局差分特征，进而提升隐含差分的可分离性。

在模型处理来自 SM4 密码算法加密数据 $\boldsymbol{X}$ 的特征编码阶段，构建了多核卷积的 Inception 多特征融合模块，通过并行的不同卷积尺度 $\{k_i\}$ 生成差分特征的多尺度映射空间 $\{\boldsymbol{z}_i\}$，其表示为

$$\boldsymbol{z}_i = k_i(\boldsymbol{X}) = k_i * \boldsymbol{X}, i = 1,\cdots,m \tag{4}$$

不同尺寸的卷积核在差分传播空间中构建一组“滤波器组”，分别聚焦局部高频（小核）与跨区相关（大核）以及全局统计特征。设真实密文与随机数据在每个分支上的均值分别为 μ_i^+、μ_i^-，分支串联后的类间均值差矢量为

$$\Delta\mu = \left[\mu_1^+ - \mu_1^-,\cdots,\mu_m^+ - \mu_m^-\right] \tag{5}$$

当某一分支 i 捕获了具有独立判别价值的差分特征时，拼接后的向量 $\boldsymbol{Z} = \left[\boldsymbol{z}_1,\boldsymbol{z}_2,\cdots,\boldsymbol{z}_m\right]$ 的 $\left\|\Delta\mu\right\|_2$ 通常显著大于任一分支单独贡献的差分特征，从而有效扩展特征子空间并提升差分特征可分性。

为了进一步增强不同特征通道之间的协同表达能力，引入 SE 注意力模块用于学习各通道差分特征的重要性权重向量 $\boldsymbol{s}$，按通道动态放大或抑制特征，其表达式为

$$\boldsymbol{s} = \sigma(W_2\delta(W_1 g)) \tag{6}$$

其中，$g = \mathrm{GAP}(\boldsymbol{Z})$ 为通道的响应统计，W_1、W_2 是 2 个全连接参数，δ 是非线性激活函数。

采用费歇尔（Fisher）判别比率 J[27]刻画特征变换后的区分能力。

$$J(\boldsymbol{Z}) = \frac{\left\|\mu_{\boldsymbol{Z}}^+ - \mu_{\boldsymbol{Z}}^-\right\|_2^2}{\mathrm{tr}\left(\sum_{w,\boldsymbol{Z}}\right)} \approx \frac{\sum_{j=1}^{C_Z}\boldsymbol{s}_j^2(\Delta\mu_j)^2}{\sum_j \boldsymbol{s}_j^2\sigma_{w,j}^2} \tag{7}$$

当权重缩放偏向具有更大 Fisher 评分的通道，即 $\dfrac{(\Delta\mu_j)^2}{\sigma_{w,j}^2}$ 更大时，可获得更优的判别效果。SE 模块通过学习映射近似于这种最优缩放，使得隐含差分特征在神经差分区分器空间中更加可分，从而提升 SM4 神经差分区分器在高轮次分析中的分类性能。

2.2 数据预处理

SM4 数据分组长度及密钥长度均为 128 bit，划分为 4 个 32 bit 输入单元，记为 $\boldsymbol{P} = (P_0,P_1,P_2,P_3)$，$P \in (\boldsymbol{Z}_2^{32})^4$。使用一组确定的差分 $\Delta = (\Delta_0,\Delta_1,\Delta_2,\Delta_3)$ 对明文 $\boldsymbol{P}$ 进行异或计算，得到 $\boldsymbol{P}' = (P_0',P_1',P_2',P_3')$。SM4 密码算法经过 S 盒加密 $\boldsymbol{P}$ 和 $\boldsymbol{P}'$ 得到密文对 $\boldsymbol{C} = (C_0,C_1,C_2,C_3)$ 和 $\boldsymbol{C}' = (C_0',C_1',C_2',C_3')$，将 $\boldsymbol{C}$ 和 $\boldsymbol{C}'$ 通过数据预处理组成训练数据格式。根据式(8)生成 Y 标签，并赋给每个数据。其中，$\mathrm{Dec}(\boldsymbol{C})$ 表示对密文 $\boldsymbol{C}$ 进行解密。

$$Y=\begin{cases}1,\mathrm{Dec}(\boldsymbol{C})\oplus\mathrm{Dec}(\boldsymbol{C}')=\Delta\\0,\boldsymbol{C}'\in\text{随机数}\end{cases}\tag{8}$$

文献[23]已经提到2种数据预处理方法，分别是输入结构$\boldsymbol{I}_1$和输入结构$\boldsymbol{I}_2$，其数据处理如式(9)和式(10)所示。

$$\boldsymbol{I}_1=(C_0,C_1,C_2,C_3,C_0',C_1',C_2',C_3')^{\mathrm{T}}\tag{9}$$

$$\boldsymbol{I}_2=(C_0,C_1,C_2,C_3,C_0\oplus C_0',C_1\oplus C_1',C_2\oplus C_2',C_3\oplus C_3')^{\mathrm{T}}\tag{10}$$

由于S盒通常表现出高度的非均匀性，而二阶差分技术能放大输入差分经过多轮传播后呈现的统计特性，形成更加明显的模式，本文提出了一种全新的输入数据结构，即高阶交叉差分密文对。该方法通过对多组密文差分进行叠加，有效削弱了线性扩散层的影响，使模型在捕获S盒的非线性特征时更加敏感。将该数据格式定义为$\boldsymbol{I}_3$，数据处理如式(11)所示。

$$\begin{aligned}\boldsymbol{I}_3=(&C_0\oplus C_0',C_1\oplus C_1',C_2\oplus C_2',C_3\oplus C_3',\\&(C_0\oplus C_0')\oplus(C_1\oplus C_1'),(C_2\oplus C_2')\oplus(C_3\oplus C_3'),\\&(C_0\oplus C_0')\oplus(C_2\oplus C_2'),(C_1\oplus C_1')\oplus(C_2\oplus C_2'))^{\mathrm{T}}\end{aligned}\tag{11}$$

值得注意的是，式(11)数据处理方式不仅包含密文对之间的部分差分，还融合了高阶差分信息，其设计理念源于多特征提取。正如Inception模块通过并行卷积提取不同尺度特征并进行融合以获得更丰富的表征，高阶交叉差分密文对也实现了从局部细节到全局模式的特征融合，从而为模型提供了更全面的非线性特征支持。$\boldsymbol{I}_3$与$\boldsymbol{I}_1$、$\boldsymbol{I}_2$数据长度相同，均为8个32 bit，不会额外增加训练时间和内存开销。式(12)为$\boldsymbol{I}_3$的矩阵形式表示。

$$\boldsymbol{I}_3=\begin{bmatrix}1&0&0&0&1&0&0&0\\0&1&0&0&0&1&0&0\\0&0&1&0&0&0&1&0\\0&0&0&1&0&0&0&1\\1&1&0&0&1&1&0&0\\0&0&1&1&0&0&1&1\\1&0&1&0&1&0&1&0\\0&1&0&1&0&1&0&1\end{bmatrix}\begin{bmatrix}C_0\\C_1\\C_2\\C_3\\C_0'\\C_1'\\C_2'\\C_3'\end{bmatrix}\tag{12}$$

2.3　模型结构

神经差分区分器模型的网络结构如图2所示，其有4个组成部分：输入层、初始卷积层、残差塔，以及预测头。

输入层由8个32 bit数据块组成，对应于3种数据预处理方法，直接传递给后续的卷积和残差模块进行特征提取。

模块1（Module 1）构建多分支一维卷积架构，通过4组并行一维卷积层（Conv1D）（卷积核尺寸分别为1、3、5，通道数为32）实现局部特征的跨尺度捕获。采用通道维度拼接策略进行多特征融合，并引入批量归一化层（BN, batch normalization）实现梯度稳定化，形成初始特征图$F_1\in R^{b\times 128}$，其中b为批尺寸。

模块2（Module 2）是一个全局平均池化操作，实现空间信息的全局汇聚，辅助模块1获取整体特征分布。

模块3（Module 3）是残差块，由2个连续的Conv1D组成，每个卷积层包含128个通道。在每个Conv1D之后，应用批量归一化层，以提高训练过程的稳定性和效率，随后是具有ReLU激活函数的激活层。残差块的核心特点在于其跳跃连接（SC, skip connection），将输入直接添加到经过非线性变换后的输出上，实现输入与其非线性变换的线性叠加。这种结构使得在堆叠多个网络层时，能够有效缓解梯度消失问题，保持模型准确率。

模块4（Module 4）是全连接层，旨在减少过拟合并提高模型性能。该模块包括2个隐藏层，每个隐藏层包含128个神经元，并配备批量归一化和ReLU激活函数。

SE注意力机制通过全局平均池化技术对特征进行尺寸压缩，从而高效整合全局上下文信息。其核心在于将生成的全局描述与原始特征逐通道相乘，动态地赋予每个通道不同的权重，显著增强了模型对特征的感知能力和区分能力，使模型能够更精准地聚焦于关键特征，从而进一步提升整体性能。

输出层由一个神经元组成，输出结果为0或1，用于二分类任务。模型使用Adam优化算法、均方误差（MSE, mean squared error）损失函数和L2正则化，式(13)为其损失函数。

$$L(y,f(x))=\frac{1}{n}\sum_{i=1}^{n}(y_i-f(x_i))^2+c\|w\|^2\tag{13}$$

其中，$f(x_i)$为网络输出，y_i为真实标签，n为训练样本数，w为网络参数，超参数惩罚因子c设为0.000 1。

2.4　数据生成

神经差分区分器利用随机数生成器分别生成数

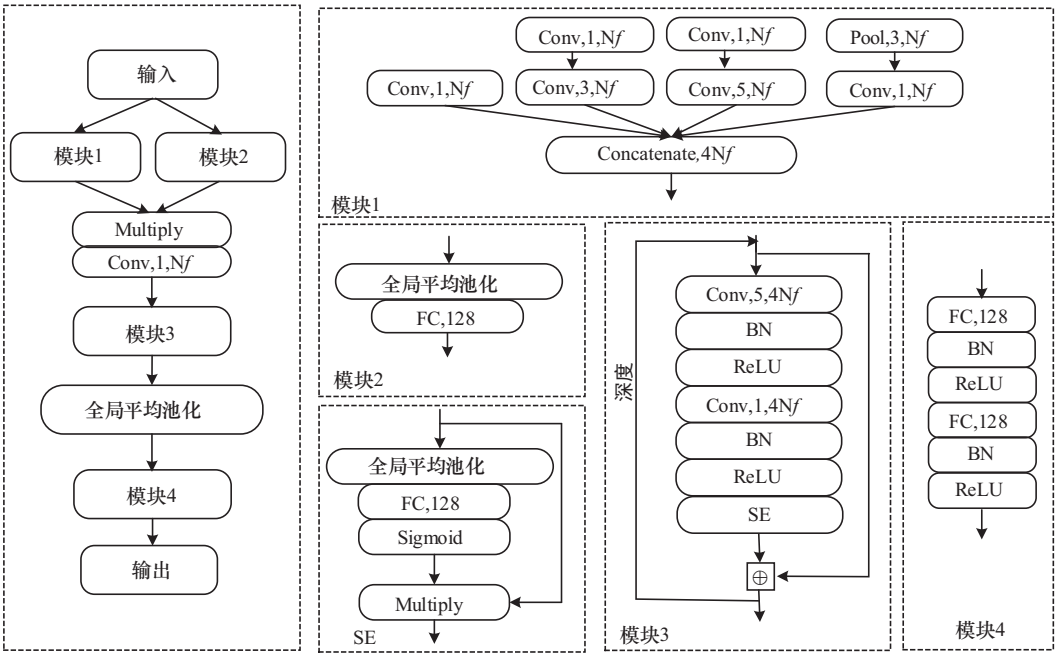

图2 神经差分区分器模型的网络结构

据集$\boldsymbol{X}$、标签Y和密钥集$\boldsymbol{K}$。其中，数据集$\boldsymbol{X}$的每个元素为128 bit，确保其尽可能随机且呈均匀分布；标签Y由0和1组成，用于标识样本的类别；而密钥集$\boldsymbol{K}$中每个密钥（key）为128 bit。3个数据集中的数据严格按照顺序一一对应。利用给定的128 bit输入差分Δ对数据集$\boldsymbol{X}$进行异或操作，生成对应的明文数据集$\boldsymbol{X}'$。将$\boldsymbol{X}$和$\boldsymbol{X}'$中的每对明文，通过密钥集K中对应的密钥进行加密计算，得到密文对$(\boldsymbol{C},\boldsymbol{C}')$。在此过程中，若标签$Y$中对应的标签为0，则将密文$\boldsymbol{C}'$替换为随机数据；若标签为1，则保持$\boldsymbol{C}'$不变。将密文对$(\boldsymbol{C},\boldsymbol{C}')$经过异或操作重组，构造出预设大小为8个32 bit的数据输入结构集合，并将该集合与标签Y一同返回。训练数据生成算法如算法1所示。

算法1 训练数据生成算法

输入 输入差分Δ和训练数据量n

输出 数据集合S

1)使用随机数生成器赋值

$$Y \leftarrow [0]\times\frac{n}{2}+[1]\times\frac{n}{2},$$

$$\boldsymbol{X} \leftarrow \text{Random}(),$$

$$\boldsymbol{K} \leftarrow \text{Random}();$$

2) 循环

3) for i=0:0:n−1

4) $\boldsymbol{X}_i' \leftarrow \boldsymbol{X}_i \oplus \Delta$；

5) $\boldsymbol{C}_i \leftarrow \text{SM4Encrypt}(\boldsymbol{X}_i,\boldsymbol{K}_i)$；

6) $\boldsymbol{C}_i' \leftarrow \text{SM4Encrypt}(\boldsymbol{X}_i',\boldsymbol{K}_i)$；

7) 如果Y_i等于0，执行$\boldsymbol{C}_i' \leftarrow \text{Random}()$；

8) end for

9) $S \leftarrow \text{BuildDataStructure}(C_{n-1},C'_{n-1})$。

2.5 模型训练和验证

模型训练即为监督学习。通过深度学习技术，模型能够自动从输入数据中提取关键特征，并学习将数据S映射到标签Y。在验证阶段，模型以输出0标记随机样本，以输出1标记密文，与真实标签一致则判为正确分类，否则判为错误分类。经过多次迭代训练后，最终保存并返回训练完成的神经网络模型，该模型应用于后续缩减轮次的密文分析工作。

模型采用贝叶斯决策框架，通过计算样本S的后验概率$P_r=(Y=1|S)$实现二元分类。依据奈曼–皮尔逊（Neyman-Pearson）准则设定决策阈值τ=0.5，当$P_r=(Y=1|S)>\tau$时判定样本源自密码算法而非随机置换。

3 实验及结果分析

本文实验均在配备GTX 1050和8 GB内存的计算机上完成，并将神经差分区分器模型应用在SM4密码算法的4~9轮测试。使用初始差分diff=(0x0, 0x1,0x2,0x4)生成2^{23}个训练集和2^{18}个测试集。

本文实验仅用数十秒生成2^{23}大小的样本，经预处理后以2^{10}批量、40个训练轮次（epoch）进行训练。学习率在小区间内循环波动，将第i个epoch的学习率l_i设置为

$$l_i = 10^{-4} + 0.001\,9 \times \frac{(\text{epoch} - 1) - i \bmod \text{epoch}}{\text{epoch} - 1} \tag{14}$$

3.1　神经差分区分器准确率

若区分器的准确率接近50%，则该区分器并不具有较好的区分性能。为了保证区分性能，当准确率超过52%时，区分器被视为有效。图3展示了4~9轮神经差分区分器在输入结构$\boldsymbol{I}_1$、$\boldsymbol{I}_2$、$\boldsymbol{I}_3$的准确率。以下实验结果均为多次重复实验后选取的表现较优的一组数据记录。在相同训练数据规模下，本文提出的多特征融合神经差分区分器在7~9轮高轮次加密场景中均取得了显著提升，准确率提升幅度均超过0.5%。其中在7轮上，相较于已有神经差分

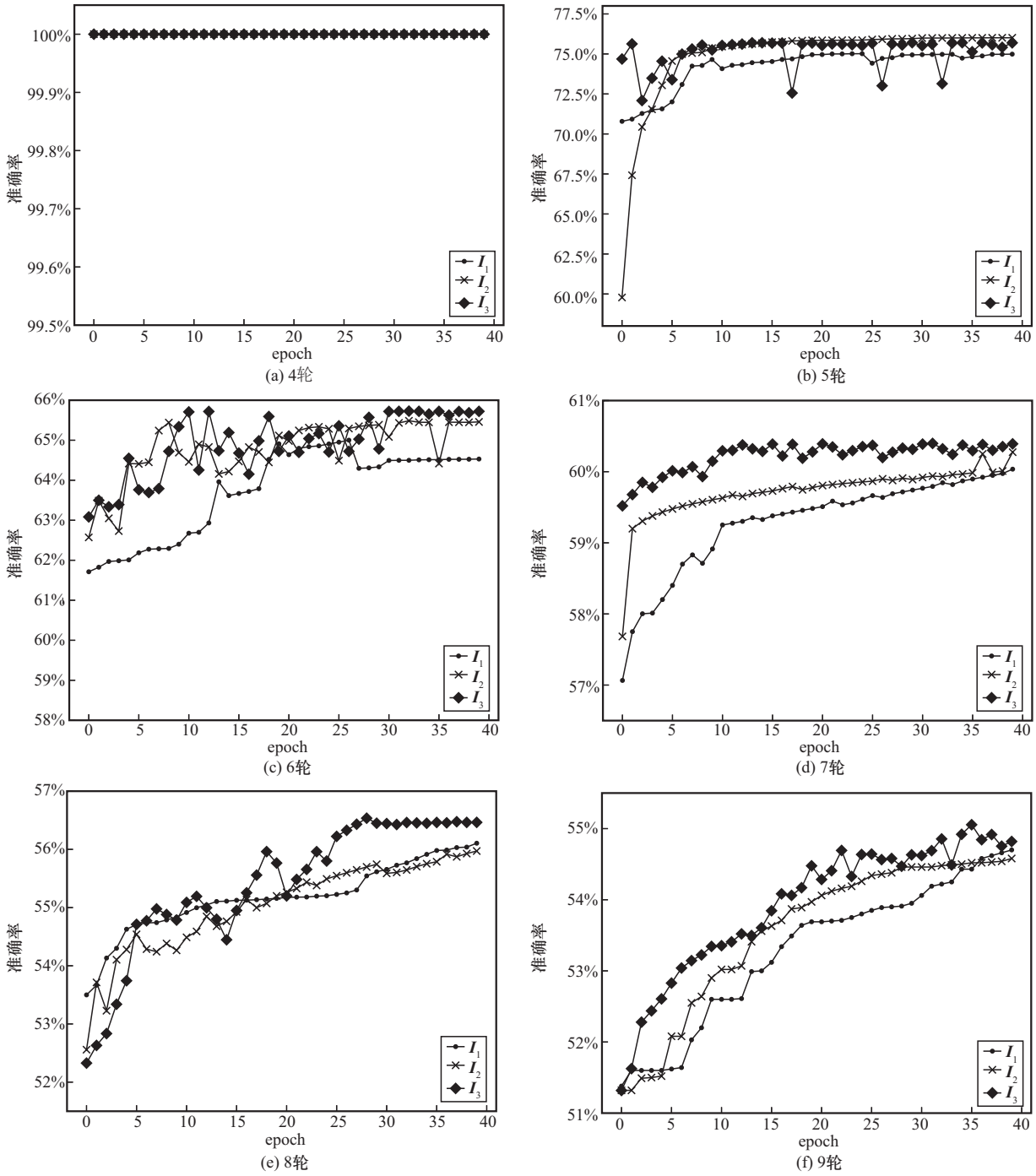

图3　SM4差分区分器的4~9轮准确率

区分器准确率，其区分准确率提升达 1.48%，这一增幅对于高轮次长分组密码的安全性分析而言具有重要意义。

根据图 3 可知，在相同实验条件下 3 种输入结构的 4~9 轮训练第一个 epoch 就得到了较高的准确率。其中 4~8 轮训练在第 1 个 epoch 就得到了有效的神经差分区分器；9 轮训练使用输入结构 $\boldsymbol{I}_3$ 在第 3 个 epoch 获得有效的神经差分区分器准确率，使用输入结构 $\boldsymbol{I}_1$ 在第 7 个 epoch 也可获得有效的神经差分区分器。区分器在 4 轮训练上使用 3 种输入结构均可获得准确率为 1 的有效神经差分区分器。在 7 轮训练上，输入结构 $\boldsymbol{I}_3$ 的训练过程中每个 epoch 的准确率均高于输入结构 $\boldsymbol{I}_1$、$\boldsymbol{I}_2$。对于神经差分区分器，在 8 轮训练上，训练周期中准确率存在波动是由于采用的是周期循环学习率，该 epoch 使用的是学习率的较大值。实验结果显示，$\boldsymbol{I}_3$ 结构较 $\boldsymbol{I}_1$、$\boldsymbol{I}_2$ 提升了准确率，尤其在 7~9 轮最为明显，准确率最高分别可达 60.39%、56.54%、55.06%。这表明基于多特征融合的神经差分区分器模型能够捕获高阶差分密文对中隐含的更加丰富的特征，为模型提供了更全面的非线性特征支持。

文献[23]采用输入结构 $\boldsymbol{I}_1$ 并未获得 9 轮符合条件的神经差分区分器，原因是文献[23]中模型在训练阶段未能准确逼近密码的差分分布，使基于结构 $\boldsymbol{I}_1$ 的特征无法有效区分密文。由图 3 可知，在经过 40 个 epoch 训练后，区分器使用输入结构 $\boldsymbol{I}_1$ 可得到准确度为 53.21% 的区分器，证明本文模型有着更强的特征提取能力，可以在结构 $\boldsymbol{I}_1$ 的密文对中捕获更多特征，所以成功获得使用结构 $\boldsymbol{I}_1$ 的神经差分区分器。另外，在训练过程中，使用结构 $\boldsymbol{I}_3$ 构建的神经差分区分器在各个 epoch 的准确率高于采用结构 $\boldsymbol{I}_1$、$\boldsymbol{I}_2$ 构建的区分器，这也证明了本文模型能够捕获更多的特征。图 4 为本文模型在不同样本量下的准确率。对于 4 轮和 5 轮神经差分区分器，2^{10} 大小的训练集即可得到有效区分；6 轮和 7 轮的训练下界为 2^{12}，在 2^{19} 后准确度差距明显；8 轮和 9 轮神经差分区分器，分别需要 2^{16} 和 2^{18} 大小的数据集才能得到超过 52% 的有效神经差分区分器。

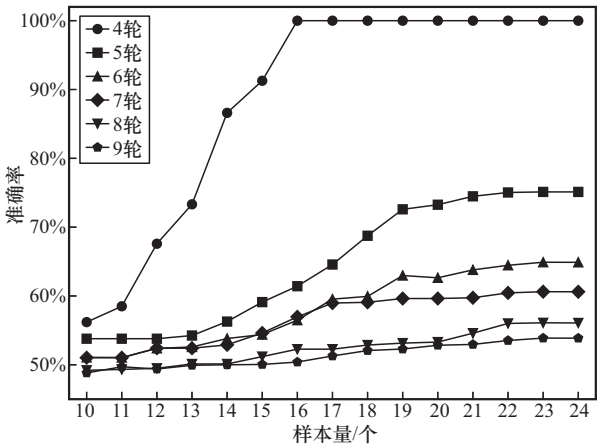

图 4 本文模型在不同样本量下的准确率

3.2 区分器对比

自 Gohr 提出利用机器学习辅助差分密码分析以来，众多学者在前人成果的基础上不断尝试和改进，相关领域已涌现出大量成果。本节开展一系列相同实验条件下自 2022 年以来每年较具有代表性的成果的对比实验，从而评估本文模型在不同轮次下的表现。

文献[23]基于残差神经网络成功构建了 9 轮神经差分区分器，并且提出了模型再学习的改进策略，在 4~9 轮上的准确率有了较大提高。本文模型在 4~9 轮上分别使用输入结构 $\boldsymbol{I}_1$、$\boldsymbol{I}_2$、$\boldsymbol{I}_3$ 进行验证，其与文献[23]中得到的准确率对比如表 1 所示。

表 1　3 种输入结构下区分器的准确率对比

轮数	本文模型			文献[23]	
	输入结构 $\boldsymbol{I}_1$	输入结构 $\boldsymbol{I}_2$	输入结构 $\boldsymbol{I}_3$	输入结构 $\boldsymbol{I}_1$	输入结构 $\boldsymbol{I}_2$
4	100.00%	100.00%	100.00%	100.00%	100.00%
5	74.24%	74.83%	75.12%	73.92%	74.98%
6	63.74%	64.72%	64.90%	63.55%	64.70%
7	59.28%	59.72%	60.62%	58.01%	59.14%
8	55.17%	55.03%	56.13%	54.22%	55.06%
9	53.21%	53.56%	53.88%	—	53.30%

观察表1使用输入结构$\boldsymbol{I}_2$、$\boldsymbol{I}_3$的实验结果可知，高阶差分能够放大输入差分经过轮传播的统计特性，削弱线性扩散层的影响，使模型更易识别S盒的非线性特征。同时，对比本文模型与文献[23]的实验结果，也可验证本文模型具有更好的特征提取能力。

为验证本文模型的性能优势，实验以文献[14]中的MLP与卷积神经网络（CNN, convolutional neural network）模型作为基准，将3类模型同步应用于SM4密码算法开展对比测试。实验结果表明，本文模型在识别准确度指标上表现出全面优越性，显著优于2类基准模型，具体数据如表2所示。

表2　本文模型与基准模型的准确率对比

轮数	本文	MLP	CNN
5	75.12%	74.16%	71.53%
6	64.90%	64.21%	62.29%
7	60.62%	58.29%	57.73%
8	56.13%	54.52%	52.90%
9	53.88%	52.89%	—

除与基准模型进行性能比对外，实验同时引入经过优化改进的代表性模型开展对比。文献[15]系统审查了当时已经提出的各类差分神经网络模型，并且对适应不同密码的神经网络模型进行了优化，探讨了模型准确性和基本分布差分之间的内在联系，对当前研究方向有重要意义，所以选择其作为对比模型之一。实验结果对比如表3所示，本文模型在5~9轮上的准确率均高于文献[15]，尤其在8轮上的准确率优势较大。

文献[16]针对攻击高轮数和大分组密码提出基于最大似然的高精度神经差分区分器构建方法，区分器准确率在7轮、8轮GIFT-128上分别达到98.8%和99.8%，在4轮ASCON密码上达到99.4%。由表2可知，文献[16]在SM4密码算法的5~7轮上仍具有较高的准确率，但在8~9轮上却不能得到有效区分器。这是由于面对SM4密码算法的高轮次加密，模型不能从中捕获足够的特征进行有效区分。文献[17]提出基于遗传算法的多差分相关密钥神经差分区分器框架，通过搜索高偏差差分并构造多差分样本，在SIMECK和SIMON密码上提升了区分准确率。文献[17]在5轮和7轮的区分任务上表现较为突出，准确率明显高于其他对比模型；尽管在其他轮次上区分性能略逊于本文模型及文献[15]，但在5~9轮上仍始终优于文献[16]。

3.3　渐进式学习策略训练

渐进式学习策略（PLS，progressive learning strategy），旨在通过分阶段引导神经网络模型逐步学习和适应复杂任务。在针对如SM4这类结构复杂的分组密码进行分析时，模型在学习过程中面临梯度消失及泛化能力下降等问题，直接训练高轮次的神经差分区分器难以取得理想效果。因此采用PLS以低轮次差分区分任务为起点训练模型，将其积累的中间轮次区分经验作为先验知识，逐步强化模型对高轮次复杂差分特征的建模能力，进而稳定训练过程并提升最终区分准确度。与文献[13]中的阶段训练（ST, staged training）类似，PLS同样通过多阶段的策略化训练，使神经差分区分器能够有效扩展至密码更高轮次的应用场景。

首先使用文献[21]的MILP自动化搜索工具识别经过5轮后输出的概率最大的差分为diff=(0xe9e17be9, 0x47e6d247, 0xae0fa9ae, 0xe93f7be9)。接着使用已经训练好的9轮神经差分区分器模型对加密6轮的数据进行区分，得到第一阶段的训练模型。然后使用第一阶段的模型训练初始差分为diff=

表3　神经区分器模型的准确率对比

轮数	本文模型	文献[15]	文献[16]	文献[17]
5	75.12%	74.90%±0.469%	74.63%±0.170%	74.98%±0.019%
6	64.90%	64.69%±0.033%	64.27%±0.716%	64.58%±3.968%
7	60.62%	58.95%±0.201%	57.28%±0.144%	59.20%±0.081%
8	56.13%	54.87%±0.127%	50.00%±0.015%	54.21%±0.248%
9	53.88%	53.51%±0.192%	50.00%±0.018%	53.22%±0.355%

(0x0,0x1,0x2,0x4)的10轮数据，重复训练2次，得到能够成功区分10轮的区分器模型。模型每一阶段训练15个epoch即可得到较为稳定的性能，使用10^{-4}的学习率，训练数据量为2^{23}，残差网络深度为1，经过参数调优验证，该配置下的模型训练效果最佳。图5为3个阶段训练随epoch变化的准确率。

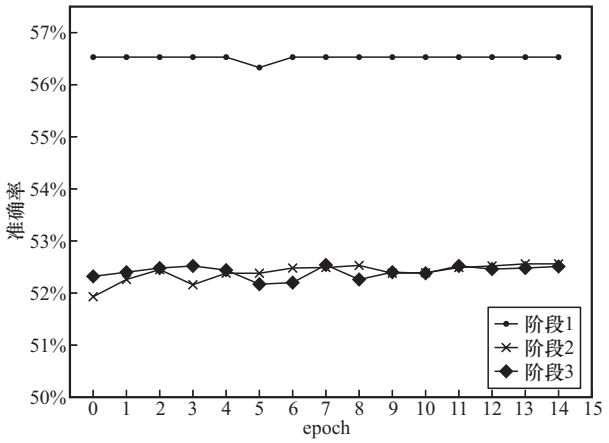

图5 3个阶段训练随epoch变化的准确率

第一阶段由于使用初始差分是5轮后的概率最大的差分，并且基于已经训练好的9轮神经区分器作为起点，因此在第一个epoch便迅速获得了有效且准确度较高的神经差分区分器。第二阶段使用第一阶段训练的模型权重，并以初始差分diff=(0x0,0x1,0x2,0x4)重新进行训练，同样在第一个epoch就得到了有效的区分器。第三阶段使用第二阶段生成的模型权重进行重复训练，以确保区分器的最终有效性。实验证明，采用渐进式学习策略，成功训练出SM4密码算法的10轮神经差分区分器，经过15个epoch训练平均准确率为52.41%，这也是目前使用深度学习技术得到的最高轮数的区分器。

此外，将输入结构$\boldsymbol{I}_1$、$\boldsymbol{I}_2$应用于模型，进行相同的渐进式学习策略操作，但均未得到有效的区分器。这进一步证明，本文模型与输入结构$\boldsymbol{I}_3$更加契合，能够捕获更多的差分特征。

模型的损失函数刻画预测与真实值的误差，是评估模型性能的重要指标。图6中展示采用渐进式学习策略得到的10轮神经差分区分器在训练集和验证集上的损失曲线，二者基本吻合，表明模型训练过程稳定且有效。

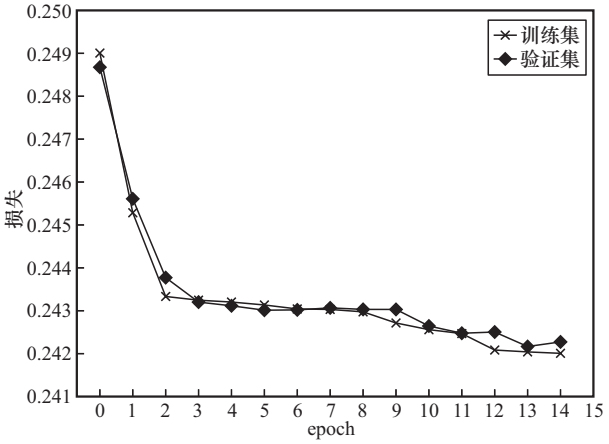

图6 区分器模型的10轮损失曲线

3.4 区分器的统计测试

敏感度衡量的是模型正确识别正类（密文）的能力，而特异度衡量的是模型正确识别负类（非密文）的能力。本文使用1 024个随机样本测试神经差分区分器的敏感度和特异度，以验证其性能。

特异度，即真阴性率（TNR, true negative rate），是正确判定为阴性的样本占所有阴性样本的比例，用以评估区分器过滤错误密文的能力，其表达式为

$$\mathrm{TNR}=\frac{2}{n}\sum_{i=1}^{n}p(Y=0|S_i)<0.5 \tag{15}$$

其中，测试样本数n=1 024。

敏感度，即真阳性率（TPR, true positive rate），是正确判定为阳性的样本占所有阳性样本的比例，其表达式为

$$\mathrm{TPR}=\frac{2}{n}\sum_{i=1}^{n}p(Y=1|S_i)>0.5 \tag{16}$$

在初始输入差分diff=(0x0,0x1,0x2,0x4)的情况下，对明文进行5~10轮加密。在选定的1 024个明文样本中，正、负样本各512个。模型采用输入结构$\boldsymbol{I}_3$，每轮测试5次并取平均值。表4为神经差分区分器的敏感度、特异度与准确率。

由表4可见，5~10轮神经差分区分器的特异度均高于敏感度和准确率，说明模型在识别非密文方面更为准确。这一现象可以归因于加密轮数的增加，使得密文的扩散性和混淆性逐步增强，导致差分特征逐渐减弱并接近随机分布。相较于正样本，非密文通常具有更明显的统计特征，因此模型更擅长“排除”非密文。

表4　神经差分区分器的敏感度、特异度与准确率

轮数	敏感度	特异度	准确率
5	51.56%	99.14%	75.12%
6	55.11%	75.55%	64.90%
7	48.40%	69.02%	60.62%
8	43.09%	62.97%	56.13%
9	50.39%	57.34%	53.88%
10	44.36%	57.70%	52.41%

图7为本文模型与基准模型的5~9轮曲线下面积（ROC）对比。本文模型在各轮次的分类性能均优于2类基准模型，其ROC整体更贴近左上角的理想区域；尤其在5~8轮测试中，本文模型的曲线下面积显著更高，表明其在真阳性识别能力与假阳性率控制间实现了更优权衡。

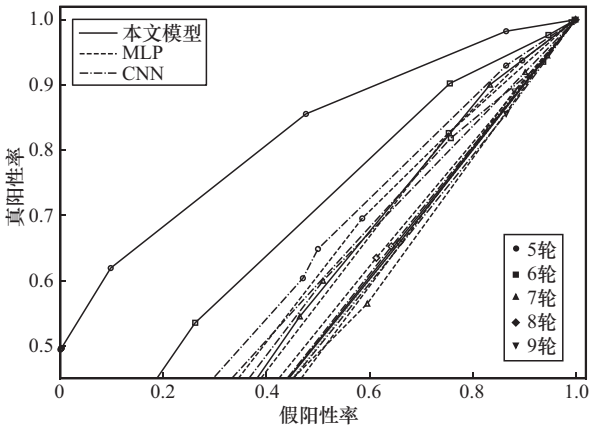

图7　本文模型与基准模型的5~9轮ROC曲线

3.5　消融实验

本节以本文提出的基于多特征融合的神经差分区分器性能为参照，通过4组消融实验验证各组件对模型性能的影响：1)禁用高阶交叉差分数据结构$\boldsymbol{I}_3$（w/o $\boldsymbol{I}_3$）；2)移除Inception模块（w/o Inception）；3)移除SE注意力机制（w/o SE）；4)取消渐进式学习策略（w/o PLS）。消融实验结果对比如表5所示。实验结果表明，禁用高阶交叉差分数据结构$\boldsymbol{I}_3$后，模型在5~9轮加密上的准确度均有所下降，其中6~8轮区分准确率降低超过1%。这说明$\boldsymbol{I}_3$能显著放大差分多轮传播后的统计特征，增强模型对隐含差分模式的表征与捕获能力，验证了高阶差分结构对增强差分可分离性的关键作用。Inception模块捕获局部与全局的复杂差分特征的能力，对模型的区分性能起着关键作用。移除该模块后，模型在各轮次的准确率均下降超过1%，其中第8轮准确率降幅达3.96%；移除SE注意力机制同样会导致模型区分准确度显著下降，虽降幅不及前者，但印证了SE注意力机制通过动态重加权突出关键差分通道，对模型建模具有重要且不可忽略的作用。Inception模块与SE注意力机制的协同作用共同支撑了模型的有效差分特征学习。模型在5~9轮差分分析中可直接获得有效区分器，而在10轮上由于高轮次差分特征更难捕获，模型未能达到预设的0.52区分阈值。通过采用渐进式学习策略，逐轮累积差分知识并增强特征表达能力，最终成功训练出10轮神经差分区分器。该结果表明差分难度随轮次呈非线性上升，渐进式学习策略对于高轮次长分组密码的安全性分析至关重要。

表5　消融实验结果对比

轮数	本文方法	w/o $\boldsymbol{I}_3$	w/o Inception	w/o SE	w/o PLS
5	75.12%	74.24%	72.05%	73.96%	—
6	64.90%	63.74%	63.01%	63.42%	—
7	60.62%	59.28%	58.05%	58.84%	—
8	56.13%	55.17%	52.17%	52.87%	—
9	53.88%	53.21%	52.82%	52.99%	—
10	52.41%	—	50.08%	50.10%	51.70%

4　结束语

本文针对现有SM4密码算法神经差分区分器存在的特征表征能力受限与攻击轮次瓶颈的问题，提出基于多特征融合与渐进式学习的区分器模型。模型使用高阶交叉差分密文对的数据预处理方法，在不增加输入数据量的前提下，使密文包含更多差异信息，同时采用Inception模块中的多核卷积层以及SE注意力机制，以提升区分器对特征的提取能力。

本文模型在5~9轮区分任务中的准确率有所提升。采用渐进式学习策略成功训练出了SM4密码算法的10轮神经差分区分器，这是目前利用深度学习技术训练SM4神经差分区分器所达到的最高轮数。通过测试模型的特异度和敏感度，并分析其训练特性，进一步验证了多特征融合技术在SM4密码算法安全性研究中的有效性，为密码分析提供了新的思路。未来将继续探索提升神经差分区分器性能的有效途径，并致力于降低对大规模训练数据的依赖。与此同时，计划将本文模型进一步扩展至

其他分组密码算法，以验证其在不同算法下的通用性与适应性。

参考文献：

[1] BIHAM E, SHAMIR A. Differential cryptanalysis of DES-like cryptosystems[J]. Journal of Cryptology, 1991, 4(1): 3-72.

[2] KNUDSEN L, WAGNER D. Integral cryptanalysis[C]//Fast Software Encryption. Berlin: Springer, 2002: 112-127.

[3] MATSUI M. Linear cryptanalysis method for DES cipher[C]//Workshop on the Theory and Application of Cryptographic Techniques. Berlin: Springer, 1994: 386-397.

[4] BIHAM E, SHAMIR A. Differential cryptanalysis of the full 16-round DES[C]//Advances in Cryptology-CRYPTO'92. Berlin: Springer, 1992: 487-496.

[5] 赵京胜, 宋梦雪, 高祥, 等. 自然语言处理中的文本表示研究[J]. 软件学报, 2022, 33(1): 102-128.
ZHAO J S, SONG M X, GAO X, et al. Research on text representation in natural language processing[J]. Journal of Software, 2022, 33(1): 102-128.

[6] RODRIGUES MOREIRA L F, MOREIRA R, TRAVENÇOLO B A N, et al. An Artificial intelligence-as-a-service architecture for deep learning model embodiment on low-cost devices: a case study of COVID-19 diagnosis[J]. Applied Soft Computing, 2023, 134: 110014.

[7] 张宇, 温光照, 米思娅, 等. 基于深度学习的二维人体姿态估计综述[J]. 软件学报, 2022(11): 4173-4191.
ZHANG Y, WEN G Z, MI S Y, et al. Overview of two-dimensional human pose estimation based on deep learning[J]. Journal of Software, 2022(11): 4173-4191.

[8] LI S H, LIN J Z, LI G Q, et al. Vehicle type detection based on deep learning in traffic scene[J]. Procedia Computer Science, 2018, 131: 564-572.

[9] YUAN Y C, SHI Y, LI C Y, et al. DeepGene: an advanced cancer type classifier based on deep learning and somatic point mutations[J]. BMC Bioinformatics, 2016, 17(17): 476.

[10] ZHANG J F, XU X L, HAN B, et al. Attacks which do not kill training make adversarial learning stronger[C]//International conference on machine learning. New York: PMLR, 2020: 11278-11287.

[11] 周玲丽, 赖剑煌. 生物特征数据安全保护技术的发展[J]. 计算机科学, 2008, 35(10): 33-38.
ZHOU L L, LAI J H. Security technology of biometric data: a survey[J]. Computer Science, 2008, 35(10): 33-38.

[12] RIVEST R L. Cryptography and machine learning[C]//International Conference on the Theory and Application of Cryptology. Berlin: Springer, 1991: 427-439.

[13] GOHR A. Improving attacks on round-reduced Speck32/64 using deep learning[C]//Advances in Cryptology-CRYPTO 2019: 39th Annual International Cryptology Conference. Berlin: Springer, 2019: 150-179.

[14] BAKSI A. Classical and physical security of symmetric key cryptographic algorithms[M]. Berlin: Springer, 2022.

[15] GOHR A, LEANDER G, NEUMANN P. An assessment of differential-neural distinguishers[J]. Cryptology ePrint Archive, 2022: 1521.

[16] YADAV T, KUMAR M. ML based improved differential distinguisher with high accuracy: application to GIFT-128 and ASCON[C]//International Conference on Security, Privacy, and Applied Cryptography Engineering. Berlin: Springer, 2024: 287-316.

[17] YUAN X, WANG Q. A Multi-differential approach to enhance related-key neural distinguishers[J]. Cryptology ePrint Archive, 2025: 697.

[18] CHENG H, DING Q. Overview of the block cipher[C]//Proceedings of the 2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control. Piscataway: IEEE Press, 2012: 1628-1631.

[19] ZHANG L, ZHANG W T, WU W L. Cryptanalysis of reduced-round SMS4 block cipher[C]//Australasian Conference on Information Security and Privacy. Berlin: Springer, 2008: 216-229.

[20] SU B Z, WU W L, ZHANG W T. Security of the SMS4 block cipher against differential cryptanalysis[J]. Journal of Computer Science and Technology, 2011, 26(1): 130-138.

[21] 潘印雪, 王高丽, 倪建强. 基于MILP寻找SM4算法的差分特征[J]. 计算机研究与发展, 2022, 59(10): 2299-2308.
PAN Y X, WANG G L, NI J Q. Finding differential characteristics of SM4 algorithm based on MILP[J]. Journal of Computer Research and Development, 2022, 59(10): 2299-2308.

[22] 王敏, 吴震, 饶金涛, 等. 针对SM4算法的约减轮故障攻击[J]. 通信学报, 2016, 37(S1): 98-103.
WANG M, WU Z, RAO J T, et al. Fault attack of reduced wheel for SM4 algorithm[J]. Journal on Communications, 2016, 37(S1): 98-103.

[23] 王慧娇, 张鑫, 韦永壮, 等. 基于深度学习的SM4密码算法新型区分器[J]. 通信学报, 2023, 44(7): 171-184.
WANG H J, ZHANG X, WEI Y Z, et al. Novel distinguisher for SM4 cipher algorithm based on deep learning[J]. Journal on Communications, 2023, 44(7): 171-184.

[24] 毛永霞, 吴文玲. 基于可分性改进分组密码SM4和FOX的积分区分器[J]. 密码学报, 2023, 10(6): 1197-1208.
MAO Y X, WU W L. Improved integral distinguishers based on division property for SM4 and FOX block ciphers[J]. Journal of Cryptologic Research, 2023, 10(6): 1197-1208.

[25] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2015: 1-9.

[26] BENAMIRA A, GERAULT D, PEYRIN T, et al. A deeper look at machine learning-based cryptanalysis[C]//Advances in Cryptology-EUROCRYPT 2021: 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2021: 805-835.

[27] BOUVEYRON C, BRUNET C. Probabilistic Fisher discriminant analysis: a robust and flexible alternative to Fisher discriminant analysis[J]. Neurocomputing, 2012, 90: 12-22.

[作者简介]

王慧娇（1976– ），女，辽宁昌图人，博士，桂林电子科技大学副教授，主要研究方向为密码学、信息安全等。

张哲（2001– ），男，河南安阳人，桂林电子科技大学硕士生，主要研究方向为人工智能、信息安全等。