

# 基于深度残差网络的人脸识别算法研究综述

方小红<sup>1</sup>, 王小生<sup>1</sup>, 刘超飞<sup>1\*</sup>, 郭桥生<sup>2</sup>

<sup>1</sup>江西理工大学理学院, 江西 赣州

<sup>2</sup>朝阳聚声泰(信丰)科技有限公司, 江西 赣州

收稿日期: 2026年3月9日; 录用日期: 2026年3月29日; 发布日期: 2026年4月9日

## 摘要

深度学习在各种各样的问题上都取得了非常好的表现, 比如物体识别、语音识别和自然语音处理。在不同类型的深度神经网络中, 深度残差网络(ResNet)得到了广泛的研究, 为人脸识别提供了新的解决方案。因此, 本文从深度残差网络的设计出发, 介绍了深度残差网络的演进及变体, 以及在人脸识别领域中常用的数据集及优化器选择、正则化方法, 之后介绍了深度残差网络及改进, 包括残差块重构、引入注意力机制、轻量化ResNet、构建ResNet系统, 在人脸识别任务中的表现。

## 关键词

深度残差网络, 人脸识别, 恒等变换, 卷积神经网络, 计算机视觉

# A Review of Research on Face Recognition Algorithms Based on Deep Residual Networks

Xiaohong Fang<sup>1</sup>, Xiaosheng Wang<sup>1</sup>, Chaofei Liu<sup>1\*</sup>, Qiaosheng Guo<sup>2</sup>

<sup>1</sup>School of Science, Jiangxi University of Science and Technology, Ganzhou Jiangxi

<sup>2</sup>Chaoyang Jushengtai (Xinfeng) Technology Co., Ltd., Ganzhou Jiangxi

Received: March 9, 2026; accepted: March 29, 2026; published: April 9, 2026

## Abstract

Deep learning has achieved remarkable performance across a wide variety of tasks, such as object recognition, speech recognition, and natural language processing. Among different types of deep neural networks, Deep Residual Networks (ResNet) have been extensively studied, providing new solutions for face recognition. Therefore, starting from the design of Deep Residual Networks, this paper

\*通讯作者。

文章引用: 方小红, 王小生, 刘超飞, 郭桥生. 基于深度残差网络的人脸识别算法研究综述[J]. 图像与信号处理, 2026, 15(2): 248-263. DOI: 10.12677/jisp.2026.152021

introduces their evolution and variants, as well as commonly used datasets, optimizer selection, and regularization methods in the field of face recognition. Subsequently, it discusses the performance of Deep Residual Networks and their improvements in face recognition tasks, such as residual block reconstruction, the incorporation of attention mechanisms, lightweight ResNet designs, and the construction of ResNet systems.

## Keywords

Deep Residual Networks, Face Recognition, Identity Transformation, Convolutional Neural Networks, Computer Vision

Copyright © 2026 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

卷积神经网络(CNN)是一种深度学习算法[1], 其在处理大规模数据时的优势日益凸显。经典的卷积神经网络从 VGG [2]、Inception、ResNet [3]发展到 DesnseNet [4], 其间也涌现了很大其他卷积神经网络, 1998 年, Yann LeCun 教授等人研究出 LeNet 体系结构[5], LeNet 模型有 7 层(不包括输入层), 可以高准确率检测支票上的数字。2012 年举办 ImageNet 竞赛, 大赛得主为 AlexNet [6], 获得广泛关注。2014 年 ImageNet 竞赛中, VGG 网络获得冠军。以上内容说明卷积神经网络的深度和宽度是改善网络性能的关键因素。但随着传统 CNN 网络深度的不断加深, 会产生梯度消失和网络退化的问题, 进而使得深层网络的性能下降。

基于传统 CNN 存在的局限性, He 等人[7]在 2016 年提出了一种深度卷积神经网络, 简称深度残差网络(ResNet), 该网络结构应用恒等映射来实现残差学习, 使不同卷积层有规律连接起来, 它解决了随着网络加深, 训练困难以及精度饱和甚至下降的问题。这使得 ResNet 在图像分类、目标检测、语义分割以及人脸识别任务中取得了出色的结果。

随着深度残差网络变得越来越受欢迎, 并在人脸识别领域取得了突破性的进展, 本文将以 ResNet 模型的发展为导向, 以基础全面的视角, 对 ResNet 核心理论基础、系列网络演进、训练方法与优化策略进行详细阐述, 并总结残差块的各种改进模型在人脸识别领域中的应用。

## 2. ResNet 核心理论基础

CNN 的基本思想是通过卷积操作和池化操作来提取数据中的空间特征, 其主要组成部分包括卷积层、池化层、全连接层和激活函数[8]。卷积层主要通过执行卷积运算提取数据特征, 是 CNN 中最核心的部分。池化层是对卷积层提取到的特征信息进行筛选, 以降低参数个数并简化网络计算复杂度。全连接层把卷积和池化后得到特征信息, 转化为更准确的模型表示。激活函数主要让网络能够捕获非线性关系并进行复杂决策。传统 CNN 可以获得图像中各个维度的特征, 并且随着网络层数的加深, 所获得的特征就更加丰富, 得到的深层信息就越多[9]-[12]。

但是在实际应用中, 深层网络并没有展现出明显优势, 例如模型训练过程会出现权重参数无法更新或损失函数无法收敛等问题, 这都是由梯度消失所造成的。另外, 增添网络层数有时会导致模型性能变差, 即准确率在加深网络层数的过程先不断提高, 然后当准确率达到饱和和状态后继续增加网络层数, 准

确率反而会降低。针对模型退化问题，一般有两种解决思想，第一种方法通过选择更好的梯度下降算法和进行更合理的权重初始化来调整求解方法或通过调整模型结构使得模型更容易优化，第二种方法是通过在网络中加入恒等映射来解决网络退化的问题。

对于传统网络很难拟合恒等映射函数，若将网络设计为：

$$H(x) = F(x) + x \tag{1}$$

如下图 1(a)所示，则可以直接将恒等映射当作网络的一部分，这就能够将问题转换为一个残差函数

$$F(x) = H(x) - x \tag{2}$$

的学习，其中若是  $F(x) = 0$ ，就形成了一个恒等映射

$$H(x) = x \tag{3}$$

网络学习的研究对象也由原始数据  $H(x)$  变成了残差  $F(x)$ 。基于该思想，Kaiming He 在 2015 年对此提出了 ResNet，它是一种改进的 CNN，ResNet 网络可以克服随着层数加深导致精度上的过拟合饱和等问题。图 1(b)表示残差模型的结构。

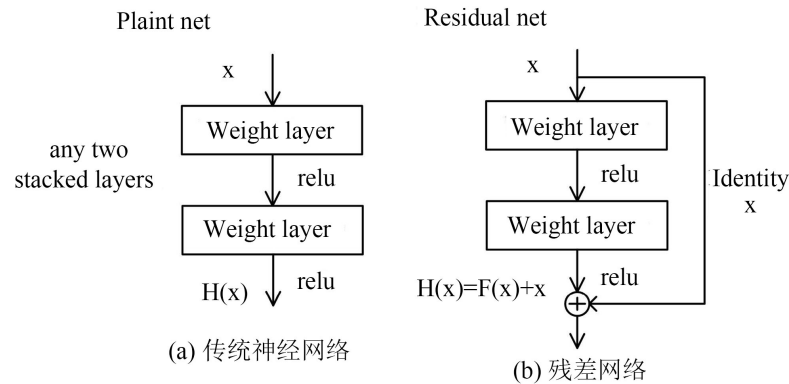


Figure 1. Schematic diagram of the residual module: (a) Traditional neural network; (b) Residual network  
图 1. 残差模块示意图：(a) 传统神经网络；(b) 残差网络

从图 1 可以看出，ResNet 方法中残差模块比传统模块图 1(a)多了 Identity 连接，ResNet 的旁支是从上一层输出直接到下一层的输入或连接到更多层，这个连接方法称为快捷(shortcut)或者跳跃连接。

将求和结果作为本层输出，如果中间层对输出结果无益，则网络会向

$$F(x) = 0 \tag{4}$$

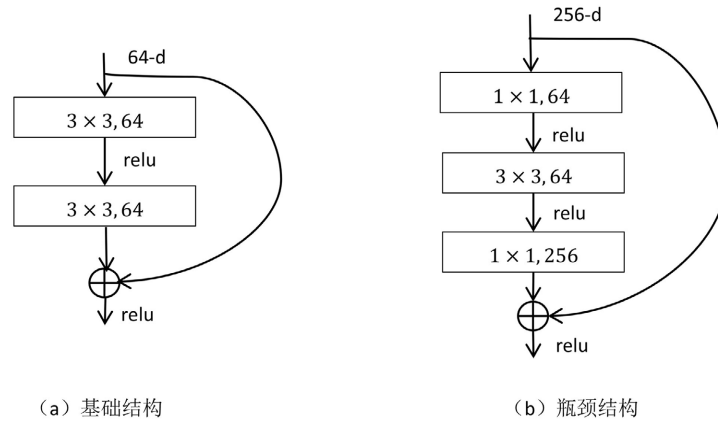
方向拟合。此外，残差模块中的权重参数更容易学习，要形成恒等映射，只需

$$F(x, w) = 0 \rightarrow w = 0 \tag{5}$$

即向权重  $w$  为 0 的方向学习。其中，残差模型中  $\oplus$  要求参与运算的  $F(x)$  与  $x$  维度相同，通常在恒等映射处加入  $1 \times 1$  卷积层对数据进行维度变换，确保输入输出可以加在一起。恒等映射块的添加使得学习目标发生变化，训练过程中不需要学习最后的输出数据，只需要学习输入数据与输出数据结果的区别，这就是残差。

为了加深神经网络的堆叠深度，在 ResNet 模型的残差块中，设计了两种特征提取结构，其主要由卷积层、批归一化、ReLU 激活函数组成。其中，基础结构由两个  $3 \times 3$  卷积块构成，在层数较小的 ResNet18 和 ResNet34 表现出优秀的特征提取能力，如图 2(a)所示；而瓶颈结构主要由 3 个卷积数分别为  $1 \times 1$ 、 $3 \times$

3、 $1 \times 1$  的卷积层堆叠而成，如图 2(b)所示，将通道数先经过  $1 \times 1$  卷积层从 256 降到 64，经过  $3 \times 3$  卷积层学习后又经过第二个  $1 \times 1$  卷积层还原到 256 的通道维度，通过这种低纬度的特征空间学习能有效地降低计算量，同时保证模型的识别性能。

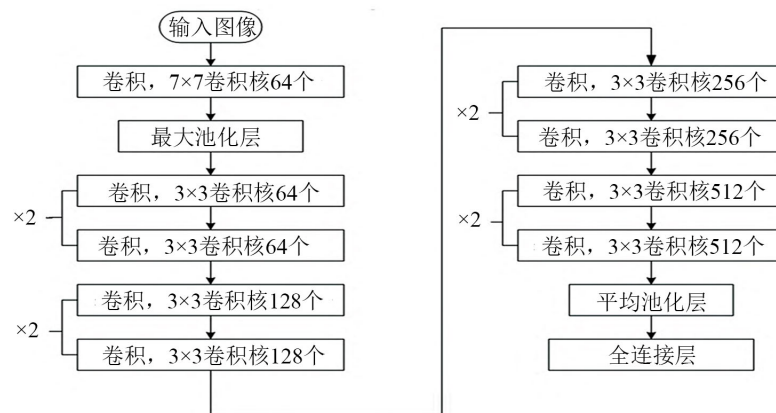


**Figure 2.** ResNet architecture diagram: (a) Basic block; (b) Bottleneck block  
**图 2.** ResNet 结构图: (a) 基础结构; (b) 瓶颈结构

### 3. ResNet 系列网络的演进

ResNet 模型提出以来，陆续出现了很多改进版本，表 1 给出了不同的 ResNet 网络的网络层数。其中常用的有 ResNet18 [13]、ResNet34 [14]、ResNet50 [15]、ResNet101 [16]、ResNet152 [17]，而这些不同网络层数的具体结构都是用 ResNet 残差块叠加而成的。下文对 ResNet18、ResNet34 与 ResNet50 这三种网络模型进行简单介绍。

ResNet18 模型由 17 个卷积层和 1 个 FC 构成。其中 17 个卷积层可分为五组，首先是  $7 \times 7 \times 64$  的卷积，卷积核大小是  $3 \times 3$ ，图 3 为 ResNet18 的结构。



**Figure 3.** Architecture diagram of ResNet18  
**图 3.** ResNet18 结构图

ResNet34 模型由 33 个卷积层和 1 个 FC 构成，其中 33 个卷积层可分为 5 组，首先是  $7 \times 7 \times 64$  的卷积，卷积核的大小分别为  $7 \times 7$  和  $3 \times 3$ ，相比普通卷积神经网络每两层间增加了跨层连接线，这就形成了残差学习结构，它的结构看上去通俗易懂，有利于进行迁移学习或应用于其他图像学习任务。如表 1 所

示, 包括 conv(卷积层)和 max pool(最大池化层), 中间使用 Relu 激活函数和批归一化层。残差模块, 包含基础残差结构。其中瓶颈结构输出的维度与输入的维度相同, 通过串联加深网络。输出模块, 将特征经过 AvgPool 平均池化层进行降采样, 输入全连接层, 由 Softmax 分类器输出对应的类别。

ResNet50 模型由 49 个卷积层和 1 个 FC 组成, 其中 49 个卷积层可分为 5 组, 卷积核大小分别为  $7 \times 7$ 、 $1 \times 1$ 、 $3 \times 3$ , 是深度卷积神经网络发展中的关键架构之一, 广泛应用于图像分类、目标检测等任务, 并常被用作性能比较的基准模型。该网络通过引入瓶颈结构(bottleneck block), 即由  $1 \times 1$ 、 $3 \times 3$ 、 $1 \times 1$  卷积组成的模块, 在控制计算成本的同时显著增加网络深度与特征表达能力。其中, 前后的  $1 \times 1$  卷积分别用于降维和升维, 有效减少了参数量与计算开销, 使得网络能够在相似计算预算下堆叠更多层。得益于残差连接的设计, 梯度得以在深层网络中稳定传播, 从而支持训练更深的模型。此外, 由于其在 ImageNet 等大型数据集上的预训练权重易于获取, ResNet50 成为迁移学习中的常用骨干网络(backbone), 广泛用于各类视觉任务的快速原型开发与模型初始化。

**Table 1.** ResNet architecture variants and configurations used

**表 1.** 采用的 ResNet 结构变体和结构

Layer name	Output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv 1	$112 \times 112$	$7 \times 7, 64, \text{stride } 2$				
		$3 \times 3, \text{max pool, stride } 2$				
conv 2_X	$56 \times 56$	$\begin{bmatrix} 3 \times 3 & 64 \\ 3 \times 3 & 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 64 \\ 3 \times 3 & 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3$
conv 3_X	$28 \times 28$	$\begin{bmatrix} 3 \times 3 & 128 \\ 3 \times 3 & 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 128 \\ 3 \times 3 & 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 8$
conv 4_X	$14 \times 14$	$\begin{bmatrix} 3 \times 3 & 256 \\ 3 \times 3 & 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 256 \\ 3 \times 3 & 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 36$
conv 5_X	$7 \times 7$	$\begin{bmatrix} 3 \times 3 & 512 \\ 3 \times 3 & 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3 & 512 \\ 3 \times 3 & 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3$
	$1 \times 1$	Average Pool, 1000-d FC, Softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

## 4. ResNet 的训练方法与优化策略

### 4.1. 常用人脸识别数据集选择与预处理

深度学习需要充分的数据支撑, 而在实际应用中, 数据往往呈现出离散型和不均衡的特点。为了使深度学习算法能够发挥效果, 必须对数据进行规范化处理。数据规范化的流程包括数据采集、数据分析和数据标注等环节, 其中数据标注是至关重要的步骤, 它将原始数据加工处理, 转化为计算机可识别的信息。在人脸识别领域中常用的数据集有以下几种:

VGG Face 223 数据集[18], 由 9131 个主题的 331 万张图像组成, 平均每个主题有 363 张图像, 覆盖各种姿势、年龄、光照、种族。

LFW (野外标记面孔)数据集[19], 它包含了超过 5700 个被试者的 13,000 多张人脸图像, 其中 1680 个被试者有两张或者更多的不同照片。

真实世界遮挡人脸(ROF) [20]数据集, 包括面部图像, 由太阳镜和口罩造成的真实上脸和下脸遮挡组成, 有 6421 张中性面部图像, 4627 张戴太阳镜的面部图像和 678 张戴口罩的面部图像。有 47 个受试者是中性的、戴面具的, 还有太阳镜图片, 114 名受试者有中性和太阳镜图片, 而 20 名受试者只有中性和蒙面图片。数据集中的身份来自一群名人和政治家, 所有的图像都来自现实生活中的场景, 在姿势和照明方面表现出实质性的变化。该数据集中的配对是通过将每个中性图像与所有戴着面具或太阳镜的受试者图像进行匹配而形成的。

UPM-GTI-Face 数据集[21], 是作为公开可用的基准, 用于评估具有挑战性的监视场景下人脸识别网络的鲁棒性, 用于评估距离和人脸面具的联合影响。该数据集包含 11 个受试者的 484 张图像, 在 3 至 30 米的 10 个注视距离上获得, 涵盖室内和室外环境, 以及 2 种口罩条件(带和不带)。对于受试者、环境和面罩条件的每一种组合, 该数据集提供了标准距离为 1 米的疑犯照片库图像, 并附有 10 个探针图像。为了确保面部验证任务的全面覆盖, 通过混合室内和室外环境精心形成成对, 同时保持蒙面和未蒙面实验的分离。因此, 对于 10 个中间距离中的每一个(每个中间距离涉及 2 个环境和 11 个受试者), 数据集包含 22 个疑犯照片库图像和 22 个探针图像。这种排列导致每个距离创建 484 对图像, 其中只有 44 对代表每个掩模条件的正确匹配(每个受试者 4 对正确匹配)。

SCface 41 数据库[22]已用于评估人脸验证的性能。该数据集包括使用五个不同质量的视频监控摄像机在不受控制的室内环境中捕获的人脸静态图像。总的来说, 该数据库包含了 130 个不同主题的面部图像。数据库中的每个受试者都由一张脸部照片图库图像和 15 张探针图像表示, 这些图像是使用 5 个不同的相机在 3 个不同的距离上拍摄的: 近距离、中距离和远距离。不同图像质量的摄像机的使用复制了现实世界的条件, 并促进了对鲁棒人脸识别算法的评估, 特别是在与执法和监视相关的场景中。通过将每个嫌疑人照片库图像与数据集中所有可用的探针图像进行比较, 建立了用于评估的配对。

The Extended Cohn-Kanade Dataset (CK)人脸表情数据集[23], 该数据集是在实验条件下获取的, 相对较为严谨可靠, 是人脸表情识别中比较常用且标准的一个公开数据集, 它包含 123 个对象的 327 个被标记的表情图片, 共分为愤怒、蔑视、厌恶、害怕、高兴、悲伤和惊讶共 7 种表情。

在某些情况下, 研究者可能会面临公开数据集不足的问题。因此, 可能需要依赖自行生成的数据集进行实验。这种情况下, 有必要清楚地说明数据的来源和生成细节, 以确保实验的可靠性和透明度。

## 4.2. 优化器设计

优化器(optimizer)就是在深度学习反向传播过程中, 指引损失函数的各个参数往正确的方向更新合适的大小, 使得更新后的各个参数让损失函数值不断逼近全局最小。卷积神经网络模型采用何种优化算法对于这个模型是否可以正常使用是非常重要的, 好的优化函数可以起到事半功倍的效果。卷积神经网络模型学习的目标是找到使目标函数的值尽可能小的参数, 这个过程称作最优化。在深度学习任务中, 网络的参数往往很大, 更加大了优化问题的复杂性。ResNet 模型训练中常用的优化器包括 SGD (小批量梯度下降算法) [24]、Adam (基于梯度下降的自适应优化算法, 能自适应调整每个参数的更新步伐) [25]。

SGD 结合了批量梯度下降算法和随机梯度下降法的优点, 以  $n$  个样本数据为一个批量执行更新。由此, 缓解了随机梯度下降法高方差的问题, SGD 的参数更新方法如下式所示:

$$\theta_t = \theta_{t-1} - \alpha \nabla_{\theta} f(\theta; x^{(i+n)}, y^{(i+n)}) \quad (6)$$

SGD 对所有参数以相同的学习率进行更新。当训练数据稀疏时, 特征的稠稀程度不一, 理想的优化

器需要对很少出现的特征赋以更大的更新步长。此外，当面对鞍点时，SGD 无法逃脱。鞍点是一个维度向上倾斜，另外一个维度向下倾斜的点，其周围所有维度的梯度均接近于 0，SGD 只考虑当前迭代的梯度更新参数，当它遇到鞍点时，当前梯度为 0，无法进行有效更新逃离鞍点。

**Table 2.** Adam algorithm

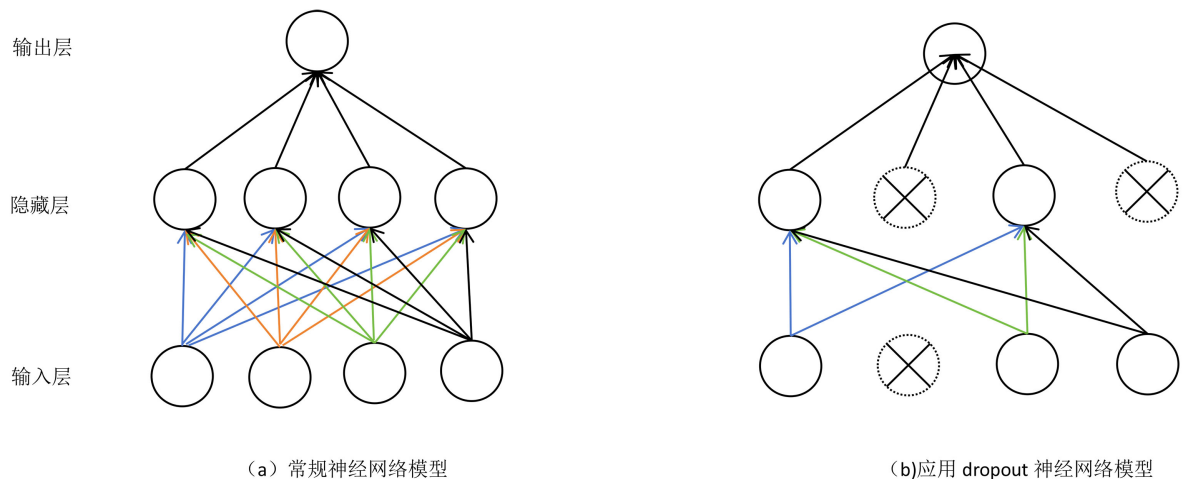
**表 2.** Adam 算法

Adam 优化算法
参数： $\eta$ ：学习率 $\beta_1, \beta_2$ ：衰减率， $0 \leq \beta_1 \leq 1$ ， $0 \leq \beta_2 \leq 1$ $f(\theta)$ ：目标函数， $\theta$ 为被优化参数 $t$ ：迭代次数 $m_0$ ：一阶向量 $v_0$ ：二阶向量 While $\theta_t$ 未收敛到最优 do $t \leftarrow t + 1$ ; $g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$ ; $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ ; $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ ; $\theta_t \leftarrow \theta_{t-1} - \eta \cdot \frac{m_t}{\sqrt{v_t + \epsilon}}$ ; end while return $\theta_t$

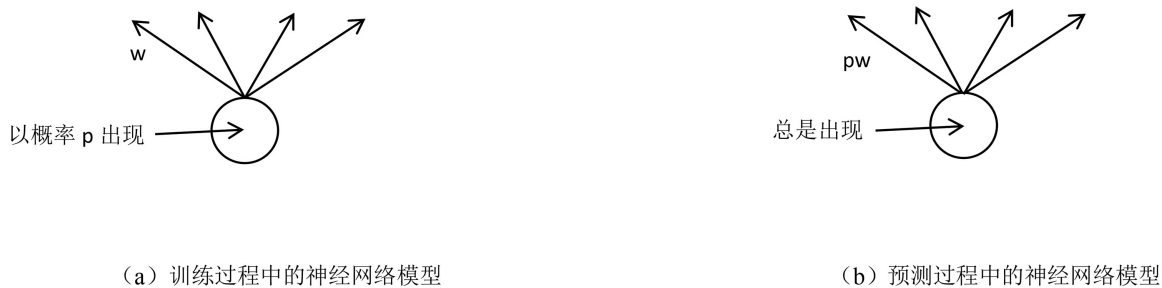
Adam 方法的详细内容见表 2， $w$  表示需要优化的参数， $v$  代表速度， $\alpha$  代表动量标量值(通常设置为 0.99)， $l$  表示误差大小， $n$  是学习率。由表 2 可知，此种方法加入了参数自适应不同学习率方法，使得这种优化算法非常快速，内存需求很少，因此它可以解决大批量的数据分析问题。

### 4.3. 正则化方法

神经网络模型参与训练过程中，往往会因为数据量的原因而导致模型的复杂度远超实际需求，还会出现过拟合，因此，Dropout [26] 这种方法被发表出来。Dropout 可以被理解作为一种剪枝操作，当把它应用在神经网络中的时候，由于网络中有众多的连接线以及单元，Dropout 可以使一些单元临时起作用，在不断重复的训练中，不同的单元都会经历这样的操作。在神经网络的训练过程中，Dropout 随机地将部分神经元的输出置零，迫使这一部分被丢弃的神经元在当前迭代中不参与计算和权重更新。通过这种方式，模型在每次训练时只能依赖部分神经元，从而减少了对特定训练数据的过度拟合，增强了模型的泛化能力。其实，从宏观上来看，Dropout 对模型结构进行了多次结合，其对应的神经网络模型结构不同，但是它们依然遵循权值共享的原则，即出现的所有这些不同的神经网络模型结构会同时共享隐含层节点的权值，这样不同的输入样本就对应不同的模型结构。图 4 展示了正常状态下的神经网络模型与采用了 Dropout 后的神经网络模型。



**Figure 4.** (a) Conventional neural network model; (b) Network model with Dropout  
**图 4.** (a) 常规神经网络模型; (b) Dropout 的网络模型



**Figure 5.** Dropout operation in the predictive model: (a) Neural network model during training; (b) Neural network model during prediction  
**图 5.** 预测模型的 Dropout 操作: (a) 训练过程中的神经网络模型; (b) 预测中的神经网络模型

概率  $p$  代表了神经元在训练时留下来的几率,  $(1-p)$  则指代不被留下的几率。测试模型阶段采用 Dropout 的示意图如图 5 所示。但是进行模型验证的时候就不会进行这些操作了, 而  $w$  和  $p$  要做乘法, 即构造  $pw$ 。测试时权重参数  $w$  需要乘  $p$  的原因如下, 以神经网络的第 1 个层为例子,  $x$  指代投入到卷积神经网络模型的数据, 当使用 Dropout 操作, 是计算出来的期望。神经元在测试阶段会一直保持活动状态, 并且使  $x \rightarrow px$ , 其中  $p$  是采用 Bernoulli 分布函数产生的 0 或 1 概率。

原来深度残差网络 ResNet 中并没有使用 Dropout, 但受 Google 提出的 GoogleNet [27] 网络模型的启发, 它虽然不是通过全连接层 (fully connected layer), 而是通过 Average Pooling 得到最终的特征, 但网络中依然使用了 Dropout, 即在全连接层前面加入了 Dropout 层, 较为容易提升网络模型的泛化能力。

## 5. ResNet 在人脸识别任务中的表现

ResNet 凭借其强大的特征提取能力和泛化性能, 已成为人脸识别领域的核心技术之一。自提出以来, 研究者们从不同的角度出发 (例如重构残差块、引入注意力机制、轻量化残差结构、构建 ResNet 系统), 对深度残差网络进行了不同方向的改进, 增强了模型的表示能力和泛化能力, 并将其应用到人脸识别领域。本章从以下五个方面, ResNet 原结构的应用、重构残差块的应用、引入注意力机制、轻量化 ResNet、构建 ResNet 系统, 来说明 ResNet 在人脸识别任务中的表现。

## 5.1. ResNet 原结构的应用

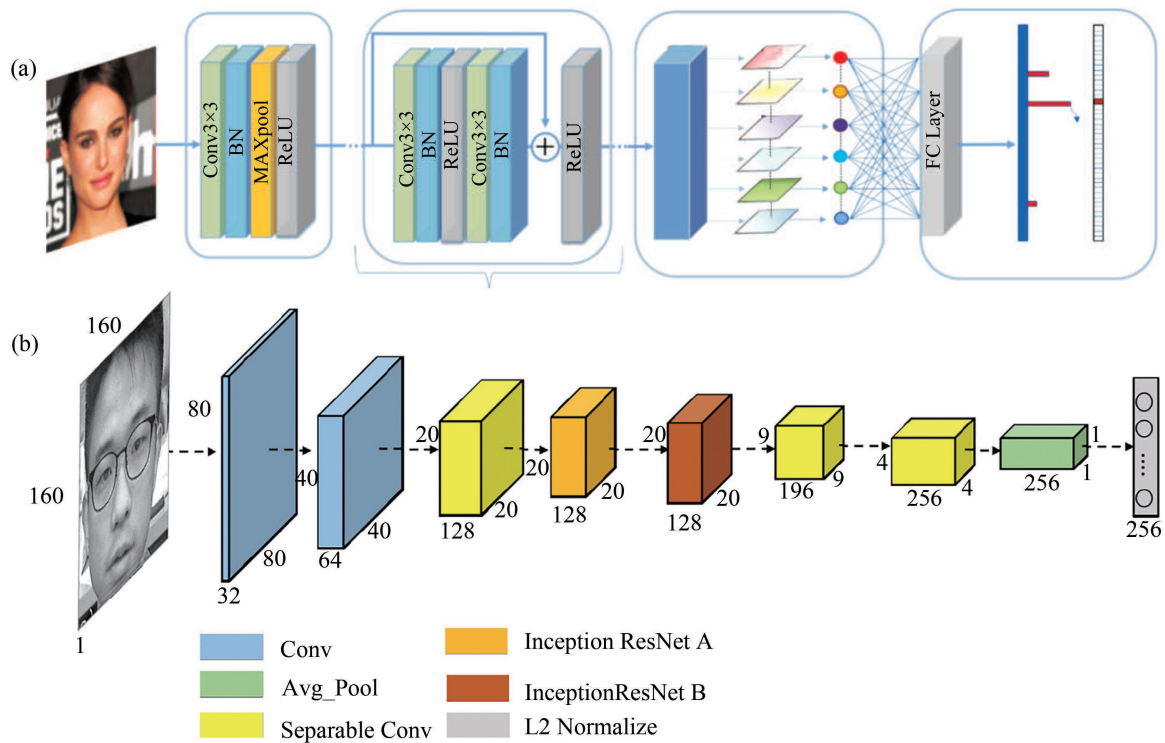
深度残差网络作为人脸识别领域的直接分类器。如: Praveen 等人[28]直接把 ResNet50 用于 FER-2013 七类情绪识别, 解决了传统模型难以兼顾特征捕捉能力与计算资源效率的问题。同时, 金力等人[29]用 ResNet50 识别 Jaffe 表情, ROC 曲线中表现效果优秀, 能够获取动态面部表情变化, 促进教学质量提高。Liu [30]用 ResNet50 加数据增强解决光照干扰, 模型展现出强鲁棒性与实用价值, 可广泛用于多领域身份验证场景。易锋等人[31]用 ResNet25 加度量学习做行人人脸识别, 完成了复杂场景下行人人脸的精准识别任务。卢官明等人[32]简单堆叠残差单元至 74 层, 靠混合数据集训练提升性能, 解决深度网络性能退化问题。Jain 等人[33]用 ResNet50 迁移学习做口罩二分类, 解决了公共场景下口罩佩戴自动检测难题。Wangean 等人[34]对 ResNet50 做参数调优(Adam, lr = 0.001)来适配小样本考勤, 显著优于传统方法, 可稳定应用于员工考勤系统。Kumar 等人[35]用 ResNet50 加 NMS 后处理适配多人脸检测, 兼顾检测效率, 适用于各类实时性要求高的应用场景。这些研究的共同特征是: 网络结构零改动, 改动的是学习率、损失函数、数据增强策略。

此外, 原始深度残差网络也可作为人脸识别系统集成的功能模块。Medjdoubi 等人[36]在 ESP32-CAM 边缘端调用 ResNet 提取特征, 接入 KNN 分类器, 完成智慧城市场景下的实时人脸识别、异常预警与远程监控任务。方国康等人[37]在 RK3399 ARM 平台上用 Mali-GPU 加速 ResNet 特征提取, 完成嵌入式场景下的实时人脸检测、追踪与识别任务。张立军等人[38]在课堂监测系统中用 ResNet 提取 128 维特征向量实现出勤统计, 用于课堂多人监测, 提高监测准确率和实时性。这些研究的共同特征是将 ResNet 融入其他传统硬件设备进行功能巩固。

随着深度残差网络的广泛应用, 其也被作为性能对比的基准基线。如 Rodrigo 等人[39]将 ResNet 与 ViT\_B32 及其他 5 类 CNN 做全方位对比, 完成人脸识别任务与验证任务的全面实验与分析。而 Komlavi 等人[40]也将深度学习(含 ResNet)与传统机器学习做跨数据集对比, 发现深度学习算法表现最优。之后, Nguyen 等人[41]将 ResNet50 与 SVM、VGG16 做多场景实测对比, 最终 ResNet50 表现最优, 准确率显著优于 SVM, 且比 VGG-16 对遮挡行为的抗干扰能力更强。而 Safwat 等人[42]将 GAN + Attention-ResNet50 的改进效果与原始 ResNet50 作对照, 改进后的 ResNet50 具备更强的抗干扰能力与泛化性能。这些研究都说明深度残差网络慢慢成为了人脸识别领域的基准模型。

## 5.2. 重构残差块的应用

研究者认为, 标准 ResNet 的残差块设计使用卷积和恒等映射结构虽然有效, 但并非最优解。于是, 研究者们开始深入到残差单元的数学结构和计算拓扑内部, 基于传统 ResNet 网络的基本结构, 进行了优化。且他们提出的这些改变无法通过简单的模块插入或层间裁剪实现, 必须重写残差块的核心。如陈希彤与卢涛[43]的 GSC-ResNet 在残差网络后端插入全局深度可分离卷积层, 采用  $3 \times 3$  小卷积核提取局部细节特征, 如图 6(a)所示, 使用全局深度可分离卷积调整特征权重, 形成了降低计算复杂度的轻量化人脸识别网络, 在提升人脸识别精度, 兼顾实时性的同时, 存在一定的设计与训练代价, 适用于对精度和实时性均有要求的复杂人脸识别场景。之后, 周婕等人[44]将 ResNet 模型的下采样操作从捷径分支移入  $3 \times 3$  卷积层, 重构了残差通路的信息流拓扑, 引入数据增强, 改进后的模型完成了愤怒、快乐、中性等七类表情的实时识别任务, 实现了计算量显著降低, 且满足实际应用要求, 适用于实时人机交互、智能监控等对模型轻量化与识别均有要求的表情识别场景。而 Deng 等人[45]提出的 FN8 则通过异质卷积核( $3 \times 5/5 \times 3$ )彻底重构了 Inception-ResNet 的内核, 如图 6(b)所示, 并引入直方图可视化、数据增强等改进, 完成了资源受限场景下的精准人脸识别任务, 可灵活部署于嵌入式门禁、车载授权等实时场景。



**Figure 6.** (a) Architecture of the global depthwise separable convolution residual network (Note: Figure sourced from [43]); (b) FN8 model (Note: Figure sourced from [45])

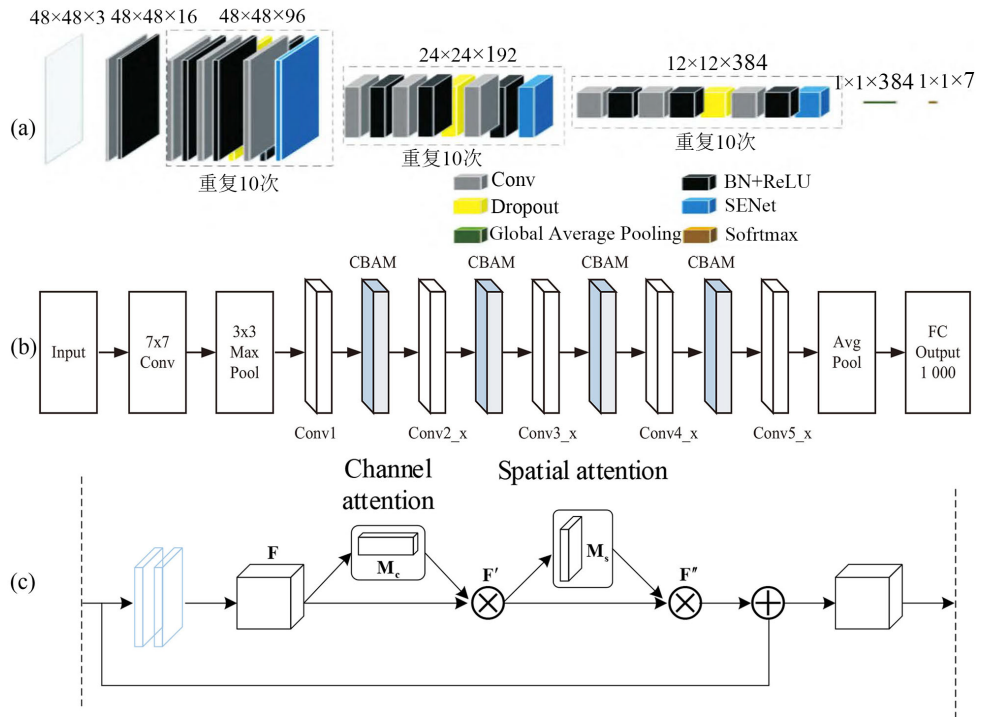
**图 6.** (a) 全局深度分离卷积残差网络结构(注: 该图来源于文献[43]); (b) FN8 模型(注: 该图来源于文献[45])

### 5.3. 引入注意力机制

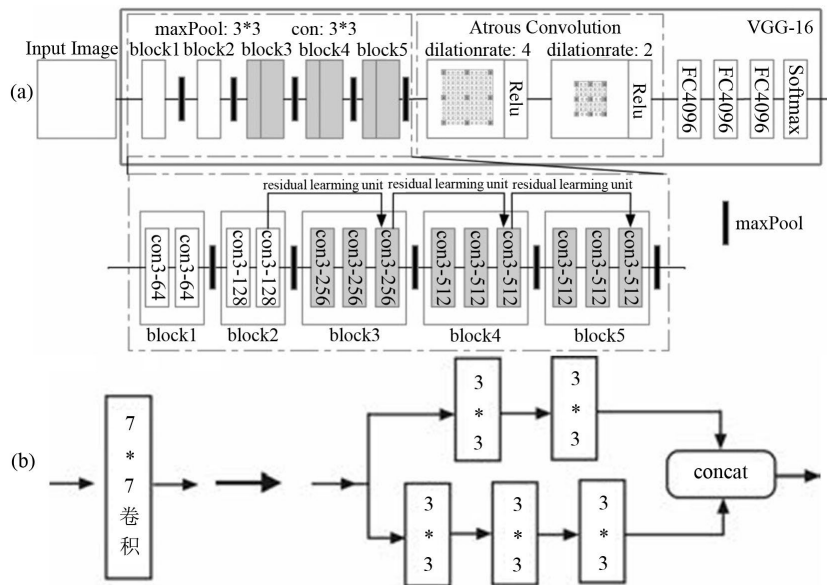
传统人脸识别通过特征提取和分类两个步骤, 但原始 ResNet 缺乏对特征重要性的差异化响应机制, 也缺乏对多尺度感受野的主动适配能力。于是, 引入注意力机制成为人脸识别领域的重要步骤。如倪锦园等人[46]的 NSENet 在宽残差模块中集成 SE 通道注意力, 如图 7(a)所示, 并引入 Dropout 正则化、学习率衰减等改进, 使得 NSENet 模型攻克了自然场景下表情识别梯度消失、过拟合及特征冗余的痛点, 改进后的模型具备更强的抗噪声与泛化能力, 但也带来一定参数计算量和调试代价, 适用于噪声、遮挡、光照变化等复杂环境下的人脸表情识别任务。而张栋昱等人[47]则更进一步地, 在 ResNet101 的 layer 间嵌入 CBAM (通道与空间双注意力), 如图 7(b)所示, 并使用 PreLU 激活函数替换 ReLU 避免梯度消失, 在 AvgPool 与 FC 层之间加入 Dropout 抑制过拟合的人脸表情识别框架, 此模型的鲁棒性与泛化能力显著优于传统模型, 适用于对准确性和鲁棒性要求较高的人机交互、情感分析等表情识别场景。之后, 石敏[48]同样在 ResNet18 中嵌入 CBAM 注意力模块, 如图 7(c)所示, 通过双分支网络聚合全局特征, 用于强化表情识别的全局纹理特征、解决表情特征单一、抗干扰能力弱的困境, 较传统单一特征方法精度显著提升, 具备更强的抗干扰能力, 这也带来了结构复杂度、计算量上升的代价, 适用于人机交互、医疗辅助等对鲁棒性和准确性要求较高的复杂表情识别场景。

此外, 另一个改变是感受野调节。如 Lai 等人[49]的 MACNN 在 VGG 骨架中插入双空洞卷积层, 融合 ResNet 残差块, 如图 8(a)所示, 专门应对微表情这种“幅度小、持续时间短”的微弱信号, 运用此模型提高了微表情识别中的特征提取能力, 收敛速度也有所提升, 实时性方面也得到提升, 但此模型也带来了一定结构设计、超参数调试与预处理开销, 适用于公共安全、心理治疗等对微表情识别精度与实时性均有较高要求的场景。而高涛等人[50]将 ResNet 输入端的  $7 \times 7$  大卷积核替换为 3 个并行  $3 \times 3$  小卷积, 如图 8(b)

所示, 用等效感受野替代单一感受野, 提高了在应对复杂表情特征、多尺度特征的提取能力, 优于传统方法与原始残差网络模型, 适用于人机交互、安全驾驶等对表情识别精度与鲁棒性要求较高的场景。



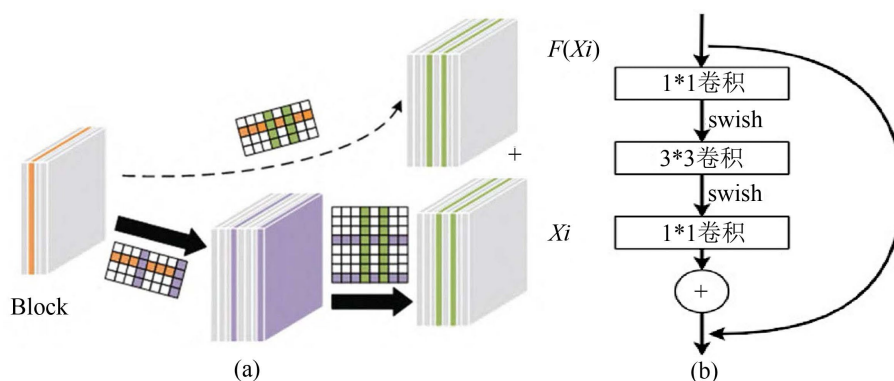
**Figure 7.** (a) Wide residual attention model (Note: Figure sourced from [46]); (b) ResNet integrated with CBAM (Note: Figure sourced from [47]); (c) CBAM-embedded residual block (Note: Figure sourced from [48])  
**图 7.** (a) 宽残差注意力模型(注: 该图来源于文献[46]); (b) ResNet 融合 CBAM(注: 该图来源于文献[47]); (c) 嵌入 CBAM 的残差块(注: 该图来源于文献[48])



**Figure 8.** (a) Architecture of MACN model (Note: Figure sourced from [49]); (b) Replacement of  $7 \times 7$  convolution with smaller convolutions (Note: Figure sourced from [50])  
**图 8.** (a) MACN 的模型结构(注: 该图来源于文献[49]); (b) 使用小卷积替代  $7 \times 7$  卷积(注: 该图来源于文献[50])

## 5.4. 轻量化 ResNet

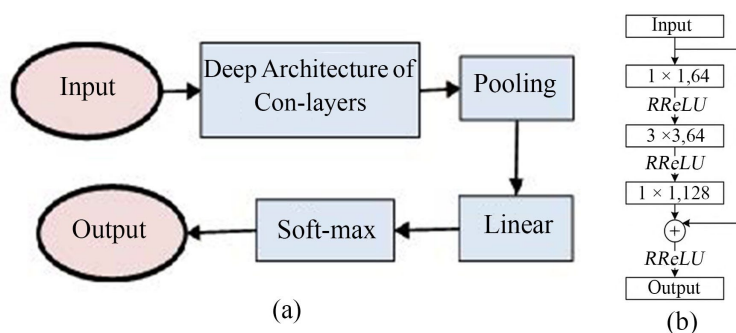
尽管卷积神经网络使遮挡人脸识别效果有所改善,但随着性能的提高会造成结构复杂、网络参数增加、耗时长的问题。因此,研究者们开始对深度残差网络进行轻量化设计。技术工具箱包括剪枝、蒸馏、瓶颈化。如刘瑞明等人[51]使用结构化剪枝,不改变残差块结构,直接修剪 ResNet50 残差块中冗余的卷积滤波器,如图 9(a)所示,修改之后的模型速度提升 36%~43%,精度仅掉 0.76 个百分点,在口罩遮挡、随机遮挡场景中均展现出精度与效率的优势。而贺怀清等人[52]提出的 DSLR (Depthwise Separable Light ResNet)融合知识蒸馏与深度可分离卷积的轻量化残差网络,采用离线的知识蒸馏迁移深层特征知识,以 ResNet50 为教师,10 层残差网络为学生,将深层特征知识迁移过去,同时集成深度可分离卷积和 SE 注意力,搭建了一个轻量级人脸识别网络,此模型能够在推理速度得到有效提升的同时识别精度基本不下降。



**Figure 9.** (a) Pruning mechanism of ResNet50 (Note: Figure sourced from [52]); (b) Improved residual block (Note: Figure sourced from [53])

**图 9.** (a) ResNet50 剪枝原理(注:该图来源于文献[52]); (b) 改进后的残差块(注:该图来源于文献[53])

之后,吴宇豪等人[53]提出的 S-ResNet 和 Obaidi 等人[54]的 A-ResNet 则选择了结构精简的改进方法,如图 9(b)所示,前者将残差模块改为  $1 \times 1 + 3 \times 3 + 1 \times 1$  的瓶颈结构,后者直接移除冗余的线性层、BN 层、Sigmoid 层,如图 10(a)所示,适用于实时人机交互、安防监控等对速度与准确性均有要求的应用场景。而苗壮等人[55]的 P-ResNet-VGG 在并行融合的框架下,也对单侧 ResNet 进行了预激活瓶颈化改造,如图 10(b)所示,使模型参数量更精简、鲁棒性更强。这些改进方法在人脸识别领域都取得了很不错的效果,解决了 ResNet 在复杂网络中内存消耗大和无法满足实时性的问题,但也带来了一定双分支结构与计算代价,适用于人机交互、智能监护等对识别精度、轻量化与泛化能力均有较高要求的表情识别场景。



**Figure 10.** (a) Architecture of A-ResNet (Note: Figure sourced from [54]); (b) Improved residual block (Note: Figure sourced from [55])

**图 10.** (a) A-ResNet 结构(注:该图来源于文献[54]); (b) 改进残差块(注:该图来源于文献[55])

## 5.5. 构建 ResNet 系统

随着 ResNet 的发展, 研究者们认为它是当前最好的特征提取器, 然后为它配备最好的外围设备, 以解决以下三个问题:

1) 修复“初次输入太差”的问题。李靖宇等人[56]用低分辨率人脸图像进行超分辨率重建, 再把修复后的图像送进改进型残差网络做特征提取, 进行图像识别, 有效解决低分辨率微小人脸识别难题。而刘静[57]用 ResNet 本身重建光照干扰图像, 通过白平衡技术补偿光照不均衡, 再利用改进型 CNN 提高区分度特征并完成身份识别, 这一改进也带来了模型复杂度和预处理代价, 适用于户外监控、自然场景身份验证等受复杂光照干扰严重的人脸识别任务。同时, Dai 等人[58]用 DCGAN 修复遮挡区域再融合 ResNet18, 实现精确人脸识别, 此方法展现出更强的遮挡鲁棒性, 但也带来了 GAN 训练复杂、串联推理耗时增加的代价, 适用于公共安全、身份验证等受口罩、局部遮挡干扰强烈的人脸识别场景。而 Istiqomah 等人[59]则用直方图均衡化与 Canny 边缘检测预处理后, 再输入迁移学习的 ResNet50, 同样也完成了复杂条件下的表情精准识别任务。

2) 特征融合型模型可以解决“单一特征不够用”的问题。如 Sunil 与 Hariprasad [60]将 XCEPTION 的深度可分离卷积特征与 ResNet50 V2 的残差特征进行特征级拼接, 整合二者在高效特征提取与深层模型上捕捉的双重优势, 完成面部情绪识别, 但模型参数增加, 计算开销上升, 适用于对特征提取全面性和识别精度要求高的场景。而马玉环等人[61]则将 ResNet36、局部二值模式(LPB)和长短期记忆网络(LSTM)融合, 实现空域与时域特征联合学习的图像序列表情识别架构, 较传统方法有显著提升。黄健[62]则通过二维离散小波变换提取人脸图像低频全局特征, 之后用改进 ResNet14 (14 个卷积层与 6 个残差单元)提取深层细节特征, 此模型整体性能优于单一特征识别方法。之后, Dwivedi 等人[63]和 Koodalsamy 等人[64]则在 ResNet 特征提取后接入 PCA 降维和 SVM 或 KNN 分类器。黄奥云等人[65]则以 13 层 CNN 加残差模块为核心, 通过卷积内部进化机制实现非线性特征提取, 嵌入式残差模块缓解梯度消失与网络退化, 结合多中心裁剪生成子模型加权融合, 完成低质人脸图像的精准识别任务。

3) 解决 Softmax 分类不够强的问题。如 Saravanan 等人[66]直接取 ResNet50 的 FC7 层特征, 输入支撑向量机(SVM)进行活体检测二分类, 完成人脸欺骗攻击的精准检测任务, 在多个数据集上均展现出优于 AlexNet 与 SVM 结合模型的检测性能。之后, Thanathamathee 等人[67]虽以 FaceNet 为核心, 但同样体现了深度特征提取与传统分类器优化, 研究者们不断对原始残差网络进行改进, 形成效果好的人脸识别技术。

## 6. 总结

许多研究者对 ResNet 进行了大量的研究并取得了一定的成果, 研究成果日益趋向成熟。因此, 本文对 ResNet 的发展背景、意义及研究领域作了简要阐述, 其中主要阐述了 ResNet 在人脸识别领域的成功应用, 包括重构残差块、引入注意力机制、轻量化 ResNet 原结构、重构 ResNet 系统的应用等方面, 对 ResNet 的在人脸识别领域的最新研究进展作了综述。

## 基金项目

国家自然科学基金(项目批准号: 12375014、11875149)、江西省青年井冈山学者奖励计划以及江西理工大学清江优秀青年人才计划。

## 参考文献

[1] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., *et al.* (2018) Recent Advances in Convolutional Neural

- Networks. *Pattern Recognition*, **77**, 354-377. <https://doi.org/10.1016/j.patcog.2017.10.013>
- [2] Alshehri, A. (2024) Skin-NeT: Skin Cancer Diagnosis Using VGG and ResNet-Based Ensemble Learning Approaches. *Traitement du Signal*, **41**, 1689-1705. <https://doi.org/10.18280/ts.410405>
- [3] Tufenkci, S. and Alagoz, B.B. (2026) Improved Classification of Star and Galaxy from Telescope by Using a Spatio-Spectral Feature ResNet Model. *Advances in Space Research*, **77**, 5425-5439. <https://doi.org/10.1016/j.asr.2025.12.027>
- [4] Rubin, J., Parvaneh, S., Rahman, A., Conroy, B. and Babaeizadeh, S. (2018) Densely Connected Convolutional Networks for Detection of Atrial Fibrillation from Short Single-Lead ECG Recordings. *Journal of Electrocardiology*, **51**, S18-S21. <https://doi.org/10.1016/j.jelectrocard.2018.08.008>
- [5] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., *et al.* (1989) Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, **1**, 541-551. <https://doi.org/10.1162/neco.1989.1.4.541>
- [6] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet Classification with Deep CNNs. *Proceedings of Conference and Workshop on Neural Information Processing Systems (NIPS) 2012*, Lake Tahoe, 3-6 December 2012, 1097-1105.
- [7] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [8] Liu, Q. and Bai, Z. (2026) 18F-FDG PET/CT Imaging Features-Driven Construction of Optimized CNN Model: For Post-Exercise Lower Limb Muscle Metabolism Assessment in Sanda Athletes. *Journal of Radiation Research and Applied Sciences*, **19**, Article ID: 102215. <https://doi.org/10.1016/j.jrras.2026.102215>
- [9] Nguyen, Q.D. and Thai, H. (2026) A Robust and Efficient CNN-Transformer Network for Crack Segmentation of High Resolution Images. *Engineering Structures*, **353**, Article ID: 122159. <https://doi.org/10.1016/j.engstruct.2026.122159>
- [10] Shi, Q., Zhang, Q., Xi, X. and Long, L. (2026) Study on Dual Early Warning of Coal Spontaneous Combustion Based on the BO-CNN-GRU Model. *Process Safety and Environmental Protection*, **208**, Article ID: 108535. <https://doi.org/10.1016/j.psep.2026.108535>
- [11] Afzal, I., Jichang, G., Siddiqui, F. and Fahad, M. (2026) CNN-CECA: Underwater Image Enhancement via CNN-Driven Nonlinear Curve Estimation and Channel-Wise Attention in Multi-Color Spaces. *Image and Vision Computing*, **167**, Article ID: 105916. <https://doi.org/10.1016/j.imavis.2026.105916>
- [12] Zhao, W., Xing, R., Wei, C., Sun, B., Jiang, T. and Zhao, Q. (2026) Improved Boundary-Aware Mask R-CNN Using Stereo Vision for Automated Rebar Inspection. *Automation in Construction*, **183**, Article ID: 106801. <https://doi.org/10.1016/j.autcon.2026.106801>
- [13] Liu, X., Wang, M., Yang, P., Wang, W., He, L. and Zhu, H. (2026) A GASF-ResNet18 Method for the Calibration of Hydrophones with Optical Technique in a Non-Free-Field Water Tank. *Measurement*, **266**, Article ID: 120404. <https://doi.org/10.1016/j.measurement.2026.120404>
- [14] Zhang, Y., Zhan, Y., Chen, H., Zhu, Y., Zhao, L. and Tian, Y. (2026) Research on Transmission Line Icing Classification and Recognition Algorithm Based on BiTex-ResNet34. *Cold Regions Science and Technology*, **245**, Article ID: 104845. <https://doi.org/10.1016/j.coldregions.2026.104845>
- [15] Salama, W.M. and Aly, M.H. (2026) Brain Tumor Segmentation and Classification: A CVAE-UNETR-ResNet50-VGG16 Hybrid Deep Learning Approach. *Alexandria Engineering Journal*, **135**, 433-449. <https://doi.org/10.1016/j.aej.2026.01.003>
- [16] Mahmood Khan, R.N., Hussain, L., Alluhaidan, A.S., Majid, A., Lone, K.J., Verdiyev, R., *et al.* (2025) COVID-19 Lung Infection Detection Using Deep Learning with Transfer Learning and ResNet101 Features Extraction and Selection. *Waves in Random and Complex Media*, **35**, 7886-7909. <https://doi.org/10.1080/17455030.2022.2091807>
- [17] K, S., Nikhat, H.F., Poonkodi, S. and R, P. (2024) A Novel Approach for Lung Cancer Segmentation and Classification Using MSF-Customized ResNet152. *IETE Journal of Research*, **71**, 1050-1062. <https://doi.org/10.1080/03772063.2024.2437546>
- [18] Cao, Q., Shen, L., Xie, W., Parkhi, O.M. and Zisserman, A. (2018) VGGFace2: A Dataset for Recognising Faces across Pose and Age. 2018 13th *IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, Xi'an, 15-19 May 2018, 67-74. <https://doi.org/10.1109/fg.2018.00020>
- [19] Huang, G.B., Mattar, M., Berg, T., *et al.* (2008) Labeled Faces in the Wild: A Database For Studying Face Recognition in Unconstrained Environments. *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition 2008*, Marseille, October 2008, 1-14.
- [20] Erakin, M.E., Demir, U. and Ekenel, H.K. (2021) On Recognizing Occluded Faces in the Wild. 2021 *International Conference of the Biometrics Special Interest Group (BIOSIG)*, Darmstadt, 15-17 September 2021, 1-5. <https://doi.org/10.1109/biosig52210.2021.9548293>

- [21] Rodrigo, M., Gonzalez-Sosa, E., Cuevas, C. and Garcia, N. (2022) UPM-GTI-Face: A Dataset for the Evaluation of the Impact of Distance and Masks in Face Detection and Recognition Systems. *2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Madrid, 29 November-2 December 2022, 1-8. <https://doi.org/10.1109/avss56176.2022.9959558>
- [22] Grgic, M., Delac, K. and Grgic, S. (2011) SCface—Surveillance Cameras Face Database. *Multimedia Tools and Applications*, **51**, 863-879. <https://doi.org/10.1007/s11042-009-0417-2>
- [23] Liliana, D.Y. (2019) Emotion Recognition from Facial Expression Using Deep Convolutional Neural Network. *Journal of Physics: Conference Series*, **1193**, Article ID: 012004. <https://doi.org/10.1088/1742-6596/1193/1/012004>
- [24] Yamazaki, M., Kasagi, A., Tabuchi, A., *et al.* (2019) Yet Another Accelerated SGD: ResNet-50 Training on ImageNet in 74.7 Seconds. arXiv: 1903.12650.
- [25] Sudewo, E.B., Kunta Biddinika, M., Umar, R. and Fadlil, A. (2025) Evaluating the Impact of Optimizer Hyperparameters on ResNet in Hanacaraka Character Recognition. *Preservation, Digital Technology & Culture*, **54**, 113-123. <https://doi.org/10.1515/pdte-2024-0061>
- [26] Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) Learning Representations by Back-Propagating Errors. *Nature*, **323**, 533-536. <https://doi.org/10.1038/323533a0>
- [27] Azida Muhammad, N., Ab Nasir, A., Ibrahim, Z. and Sabri, N. (2018) Evaluation of CNN, AlexNet and GoogleNet for Fruit Recognition. *Indonesian Journal of Electrical Engineering and Computer Science*, **12**, 468-475. <https://doi.org/10.11591/ijeecs.v12.i2.pp468-475>
- [28] Praveen, T.N.V.S., Sivathmika, D., Jahnavi, G. and Bolledu, J. (2023) An In-Depth Exploration of ResNet-50 for Complex Emotion Recognition to Unraveling Emotional States. *2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, Gharuan, 5-6 May 2023, 1-5. <https://doi.org/10.1109/incacct57535.2023.10141774>
- [29] 金力, 金正贤, 卢海妹, 等. 基于 ResNet 算法的课堂教学效果评价模型[J]. 兰州文理学院学报(自然科学版), 2021, 35(6): 62-66.
- [30] Liu, J. (2024) Face Recognition Technology Based on ResNet-50. *Applied and Computational Engineering*, **39**, 160-165. <https://doi.org/10.54254/2755-2721/39/20230593>
- [31] 易锋, 胡馨莹. 基于深度残差网络的行人人脸识别算法研究[J]. 电脑知识与技术, 2018, 14(23): 233-235.
- [32] 卢官明, 朱海锐, 郝强, 等. 基于深度残差网络的人脸表情识别[J]. 数据采集与处理, 2019, 34(1): 50-57.
- [33] Jain, A., Moparthy, N.R., Swathi, A., Sharma, Y.K., Mittal, N., Alhussen, A., *et al.* (2024) Deep Learning-Based Mask Identification System Using ResNet Transfer Learning Architecture. *Computer Systems Science and Engineering*, **48**, 341-362. <https://doi.org/10.32604/csse.2023.036973>
- [34] Wangean, D.A., Pangestu, G., Setyawan, S., Maulana, F.I., Gunawan, E.P. and Huda, C. (2024) The Implementation of ResNet-50 Architecture for Face Recognition Algorithm in Attendance System. *AIP Conference Proceedings*, **2927**, Article ID: 060014. <https://doi.org/10.1063/5.0205236>
- [35] Kumar, M., Gulhane, M., Kumar, S., Sharma, H., Verma, R. and Verma, D. (2023) Improved Multi-Face Detection with ResNet for Real-World Applications. *2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART)*, Moradabad, 22-23 December 2023, 43-49. <https://doi.org/10.1109/smart59791.2023.10428518>
- [36] Medjdoubi, A., Meddeber, M. and Yahyaoui, K. (2024) Smart City Surveillance: Edge Technology Face Recognition Robot Deep Learning Based. *International Journal of Engineering*, **37**, 25-36. <https://doi.org/10.5829/ije.2024.37.01a.03>
- [37] 方国康, 李俊, 王垚儒. 基于深度学习的 ARM 平台实时人脸识别[J]. 计算机应用, 2019, 39(8): 2217-2222.
- [38] 张立军, 曹江涛, 姬晓飞, 等. 基于深度学习的学生课堂学习状态监测系统设计[J]. 电子测量与仪器学报, 2024, 38(4): 37-45.
- [39] Rodrigo, M., Cuevas, C. and García, N. (2024) Comprehensive Comparison between Vision Transformers and Convolutional Neural Networks for Face Recognition Tasks. *Scientific Reports*, **14**, Article No. 21392. <https://doi.org/10.1038/s41598-024-72254-w>
- [40] Komlavi, A.A., Chaibou, K. and Naroua, H. (2024) Comparative Study of Machine Learning Algorithms for Face Recognition. *Revue Africaine de Recherche En Informatique et Mathématiques Appliquées*, **40**, 1-26.
- [41] Nguyen, T.V. and Chu, T.D. (2023) Comparative Study on the Performance of Face Recognition Algorithms. *Eureka: Physics and Engineering*, No. 4, 120-132. <https://doi.org/10.21303/2461-4262.2023.002831>
- [42] Safwat, S., Mahmoud, A., Eldesouky Fattouh, I. and Ali, F. (2024) Hybrid Deep Learning Model Based on GAN and RESNET for Detecting Fake Faces. *IEEE Access*, **12**, 86391-86402. <https://doi.org/10.1109/access.2024.3416910>
- [43] 陈希彤, 卢涛. 基于全局深度分离卷积残差网络的高效人脸识别算法[J]. 武汉工程大学学报, 2019, 41(3): 276-282.

- [44] 周婕, 马明栋. 基于改进的 ResNet 网络的人脸表情识别[J]. 计算机技术与发展, 2022, 32(1): 25-29.
- [45] Deng, Z., Chiang, H., Kang, L. and Li, H. (2023) A Lightweight Deep Learning Model for Real-Time Face Recognition. *IET Image Processing*, **17**, 3869-3883. <https://doi.org/10.1049/ipr2.12903>
- [46] 倪锦园, 张建勋, 张馨月. 深度宽残差网络注意力机制的人脸表情识别[J]. 重庆理工大学学报(自然科学), 2023, 37(1): 177-185.
- [47] 张栋昱, 赵磊. 融合注意力机制改进 ResNet 的人脸表情识别[J]. 计算机技术与发展, 2023, 33(5): 130-137.
- [48] 石敏. 基于深度残差网络与人脸关键点的表情识别[J]. 软件导刊, 2022, 21(4): 226-232.
- [49] Lai, Z., Chen, R., Jia, J. and Qian, Y. (2023) Real-Time Micro-Expression Recognition Based on ResNet and Atrous Convolutions. *Journal of Ambient Intelligence and Humanized Computing*, **14**, 15215-15226. <https://doi.org/10.1007/s12652-020-01779-5>
- [50] 高涛, 邵倩, 张亚南, 等. 基于深度残差网络的人脸表情识别研究[J]. 电子设计工程, 2020, 28(23): 101-104.
- [51] 刘瑞明, 徐春融, 周韬, 等. 基于深度残差网络模型压缩的局部遮挡人脸识别[J]. 智能计算机与应用, 2023, 13(11): 244-249.
- [52] 贺怀清, 闫建青, 惠康华. 基于深度残差网络的轻量级人脸识别方法[J]. 计算机应用, 2022, 42(7): 2030-2036.
- [53] 吴宇豪, 陈晓辉. 基于改进的 ResNet 的人脸表情识别系统[J]. 信息通信, 2020(7): 37-39.
- [54] Obaid, A.M., Shawkat, A.S. and Abdulhussein, N.S. (2023) Exploring the Potential of A-ResNet in Person-Independent Face Recognition and Classification. *International Journal of Advanced Network, Monitoring and Controls*, **8**, 12-19. <https://doi.org/10.2478/ijanmc-2023-0052>
- [55] 苗壮, 程卫月, 林克正, 等. 融合并行网络特征的人脸表情识别算法[J]. 哈尔滨理工大学学报, 2022, 27(6): 95-102.
- [56] 李靖宇, 程卫月, 李子翔, 等. 超分辨率重建的微小人脸识别算法[J]. 哈尔滨理工大学学报, 2022, 27(3): 52-58.
- [57] 刘静. 基于深度残差网络的光照干扰人脸识别方法[J]. 通化师范学院学报, 2024, 45(10): 42-47.
- [58] Dai, C. and Zeng, X. (2024) Occluded Face Recognition Network Based on DCGAN and ResNet. *Procedia Computer Science*, **243**, 724-733. <https://doi.org/10.1016/j.procs.2024.09.087>
- [59] Istiqomah, A.A., Sari, C.A., Susanto, A. and Rachmawanto, E.H. (2024) Facial Expression Recognition Using Convolutional Neural Networks with Transfer Learning ResNet-50. *Journal of Applied Informatics and Computing*, **8**, 257-264. <https://doi.org/10.30871/jaic.v8i2.8329>
- [60] Sunil, M.P. and Hariprasad, S.A. (2023) Facial Emotion Recognition Using a Modified Deep Convolutional Neural Network Based on the Concatenation of XCEPTION and RESNET50 V2. *International Journal of Electrical and Electronics Engineering*, **10**, 94-105. <https://doi.org/10.14445/23488379/ijeee-v10i6p110>
- [61] 马玉环, 张瑞军, 武晨, 等. 深度残差网络和 LSTM 结合的图像序列表情识别[J]. 重庆邮电大学学报(自然科学版), 2020, 32(5): 874-883.
- [62] 黄健. 基于深度学习与二维离散小波分解特征相融合的 adaboost 人脸识别模型[J]. 软件工程, 2020, 23(2): 43-46.
- [63] Dwivedi, R.K. and Kumar, D. (2023) ResNet and PCA-Based Deep Learning Scheme for Efficient Face Recognition. *International Journal of Intelligent Information Technologies*, **19**, 1-20. <https://doi.org/10.4018/ijit.329957>
- [64] Koodalsamy, B., Veerayan, M.B. and Narayanasamy, V. (2023) Face Recognition Using Deep Learning. *E3S Web of Conferences*, **387**, Article ID: 05001. <https://doi.org/10.1051/e3sconf/202338705001>
- [65] 黄奥云, 程宾洋, 孙家炜. 基于卷积内部进化机制与特征融合的人脸识别算法的研究[J]. 现代计算机(专业版), 2019, 25(10): 75-80.
- [66] Saravanan, P., Pandey, A., Joshi, K., Rondon, R., Narasimharao, J. and Imran, A.A. (2023) Using Machine Learning Principles, the Classification Method for Face Spoof Detection in Artificial Neural Networks. *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, 12-13 May 2023, 2784-2788. <https://doi.org/10.1109/icacite57410.2023.10182551>
- [67] Thanathamath, P., Sawangarreerak, S., Kongkla, P. and Mohd Nizam, D.N. (2023) An Optimized Machine Learning and Deep Learning Framework for Facial and Masked Facial Recognition. *Emerging Science Journal*, **7**, 1173-1187. <https://doi.org/10.28991/esj-2023-07-04-010>