

# 基于神经网络与智能手机应用的3种眼部疾病检测

卜雪奎

长江大学电子信息与电气工程学院，湖北 荆州

收稿日期：2024年7月4日；录用日期：2024年8月10日；发布日期：2024年8月30日

---

## 摘要

眼部疾病在初期可能被忽视，导致很少有人及时就医检查，从而延误疾病的诊断与治疗。为应对这一问题，将眼部疾病检测系统部署到智能手机中，使人们随时能够通过拍照来进行眼部检查，从而及早发现病情并防止其恶化。本研究旨在利用眼部特征数据集训练神经网络模型，并将其部署到智能手机设备上，以实现对白内障、葡萄膜炎、甲亢突眼这三种眼部疾病以及正常眼睛图像的识别分类。本文通过将深度学习技术与智能手机相结合，提出了针对眼部疾病检测的模型MobileEDT。该模型利用卷积神经网络提取病灶特征，同时结合Transformer的自注意力机制，实现了模型的轻量化。实验结果表明，所提出的眼部疾病检测模型在准确性(Accuracy为0.999)和效率方面均表现出较高性能。

---

## 关键词

神经网络，智能手机，眼部疾病，Transformer

---

# Detection of Three Eye Diseases Based on Neural Networks and Smartphone Applications

Xuekui Bu

School of Electronic Information and Electrical Engineering, Yangtze University, Jingzhou Hubei

Received: Jul. 4<sup>th</sup>, 2024; accepted: Aug. 10<sup>th</sup>, 2024; published: Aug. 30<sup>th</sup>, 2024

---

## Abstract

Eye diseases may be overlooked in their early stages, resulting in few people seeking timely med-

ical checkups, which delays the diagnosis and treatment of the disease. To counteract this problem, eye disease detection systems are deployed into smartphones to enable people to perform eye examinations by taking pictures at any time, thus detecting the condition early and preventing its deterioration. The aim of this study is to train a neural network model using an eye feature dataset and deploy it to a smartphone device for recognizing and classifying three eye diseases, namely cataract, uveitis, and eye protrusion due to hyperthyroidism, as well as normal eye images. In this paper, we propose MobileEDT, a model for eye disease detection, by combining deep learning techniques with smartphones. The model utilizes convolutional neural networks to extract lesion features, while incorporating Transformer's self-attention mechanism to achieve a lightweight model. Experimental results show that the proposed eye disease detection model exhibits high performance in terms of both accuracy (Accuracy of 0.999) and efficiency.

## Keywords

Neural Network, Smartphones, Eye Diseases, Transformer

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

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

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



Open Access

## 1. 引言

近年来，眼部疾病的发病率，尤其是白内障、葡萄膜炎以及甲亢突眼的病例，显著上升[1]，成为威胁人类健康的重大问题[2]。白内障是一种常见的无痛性眼病，表现为眼内晶体的混浊[3]；葡萄膜炎，亦称色素膜炎，涉及虹膜、睫状体和脉络膜的炎症[4]；而甲亢突眼则是甲状腺自身免疫损伤引发的眼眶组织病变，包括眼外肌和眼球后方脂肪[5]。

及时发现并治疗眼部疾病对于预防病情恶化至关重要。然而，传统的医院眼科检查往往耗时且不便，导致许多人在出现眼部症状时选择忽视，错失治疗良机。幸运的是，随着人工智能技术的飞速发展，医疗领域迎来了前所未有的便捷。深度学习技术的引入为眼部疾病的诊断开辟了新的研究方向。

以卷积神经网络为代表的传统人工智能方法，通过多层次卷积处理，能够有效地提取图像中的病灶信息，并通过全连接层输出分类结果[6]。这种方法不仅具备强大的特征提取能力，还拥有出色的泛化效果[7]。尽管如此，当前多数神经网络算法模型仍依赖高性能服务器或电脑运行，这在一定程度上限制了其在眼部疾病检查中的实时性。

针对这一挑战，本研究开发了一种名为 MobileEDT 的轻量级眼部疾病识别模型。该模型专为智能手机设计，能够在计算资源受限的环境下实现高精度和低复杂度的需求[8]。配备此模型的智能手机应用程序能够根据用户拍摄的眼部照片，快速判断是否存在眼疾及其类型，从而为患者赢得宝贵的治疗时间。

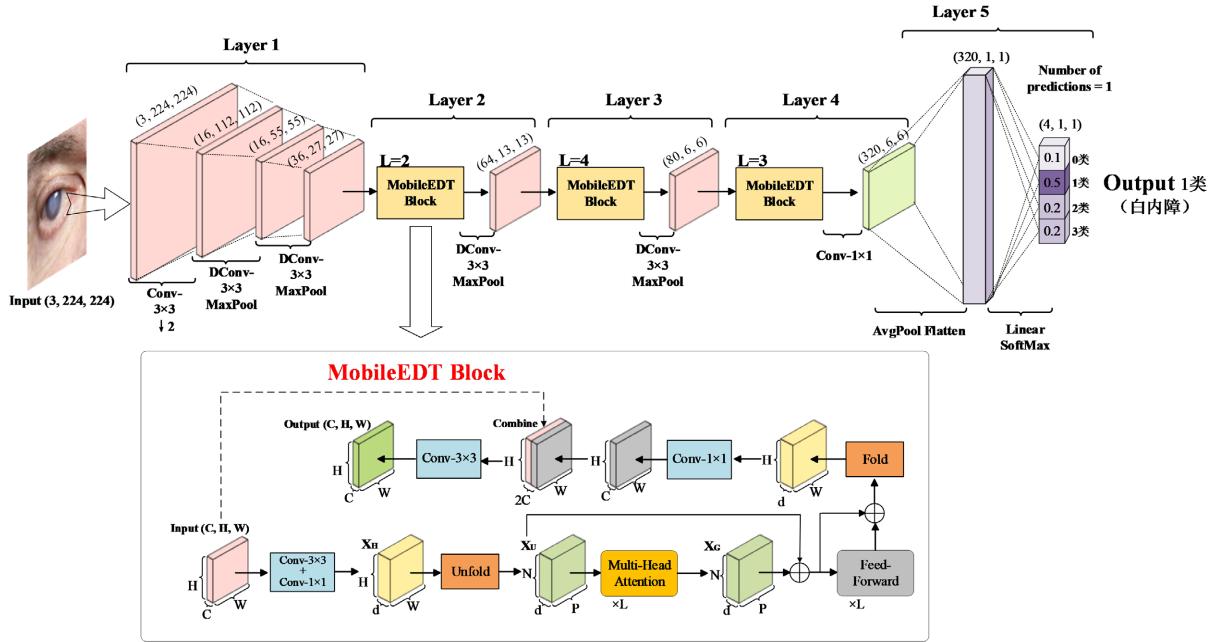
本文的剩余部分安排如下：第二节详细阐述了提出的分类模型 MobileEDT；第三节展示了 MobileEDT 与现有轻量级模型在智能手机上的综合对比实验；最后，第四节总结了研究成果以及对未来工作的展望。

## 2. MobileEDT 模型

### 2.1. MobileEDT 的网络结构

MobileEDT 模型由三个部分组成：卷积部分、MobileEDT Block 部分和分类部分，其网络结构示意图如图 1 所示。在这个分类体系中 0 类代表健康眼睛，1 类代表白内障，2 类代表葡萄膜炎，而 3 类代表

甲状腺眼。



**Figure 1.** Schematic diagram of the network structure of MobileEDT

**图 1.** MobileEDT 的网络结构示意图

图中 L 是指在每个 MobileEDT Block 中多头注意力机制和前馈网络的次数。MobileEDT 网络的每层的具体输入输出如表 1 所示。

**Table 1.** Inputs and outputs per layer of the MobileEDT deep learning model

**表 1.** MobileEDT 深度学习模型的每层输入输出

| 层       | 通道数 × 输入尺寸                | 通道数 × 输出尺寸               |
|---------|---------------------------|--------------------------|
| Layer 1 | $3 \times 224 \times 224$ | $36 \times 27 \times 27$ |
| Layer 2 | $36 \times 27 \times 27$  | $64 \times 13 \times 13$ |
| Layer 3 | $64 \times 13 \times 13$  | $80 \times 6 \times 6$   |
| Layer 4 | $80 \times 6 \times 6$    | $320 \times 6 \times 6$  |
| Layer 5 | $320 \times 6 \times 6$   | $4 \times 1 \times 1$    |

## 2.2. 卷积部分

模型使用步长为 2 的  $3 \times 3$  标准卷积用来提取输入张量的局部特征；为了提高模型效率，降低计算量和复杂度，使用步长为 1 的  $3 \times 3$  深度卷积[9]继续提炼张量特征，并在之后进行步长为 2 的最大池化，进行特征选择[10]；由于高纬度的张量经过 SiLU 激活函数之后会损失更少信息，所以步长为 1 的  $1 \times 1$  卷积是为了将张量映射到更高的维度。

## 2.3. MobileEDT Block 部分

在 MobileEDT Block 部分，模型通过较少的计算量得到输入张量的全局信息[11]。该部分可以大体分

为两层。第一层是多头自注意力机制(Multi-Head Attention)，第二层是前馈神经网络(Feed-Forward Network)。两层之间采用残差连接，然后进行层归一化。具体而言，每一层的输出为  $\text{LayerNorm}(x + \text{Sublayer}(x))$ ，其中  $\text{Sublayer}(x)$  是所在层的实现函数，MobileEDT Block 具体设计方法说明如下：

- 卷积：

MobileEDT Block 中的卷积操作旨在变换张量的大小，其中卷积核移动步长均为 1。

- Unfold：

Unfold 操作使得模型能够学习带有空间归纳偏差的全局信息，根据输入张量中元素的位置关系将  $X_H \in R^{d \times H \times W}$  均分为  $N$  份，每一份含有的元素  $X \in R^{d \times l \times l}$  个数为  $P$ ，所以按每份对应位置的元素 Unfold 之后得到  $X_U \in R^{d \times N \times P}$ ，其中  $d$  为张量的空间维度。

- Multi-Head Attention [12]：

将调整好位置的  $X_U \in R^{d \times N \times P}$  按列( $P$  列)送入多头自注意力机制中来提取全局信息，对于每一列的输入张量  $X_{U\text{-column}} \in R^{d \times N \times 1}$ ，根据公式(1)得到输出张量  $X_{G\text{-column}} \in R^{d \times N \times 1}$ 。

$$X_{G\text{-column}} = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (1)$$

其中  $Q, K, V$  是输入张量  $X_{U\text{-column}} \in R^{d \times N \times 1}$  经过矩阵变换得到的。 $P$  列的输出合在一起就是  $X_G \in R^{d \times N \times P}$ ， $X_G$  包括全局特征和局部特征。

- Feed-Forward Network：

位置偏置的前馈神经网络由 SiLU 激活函数层和一个线性层组成，输入为  $X_G \in R^{d \times N \times P}$ ，根据公式(2)输出张量的维度为  $X \in R^{d \times N \times P}$ 。

$$X \in R^{d \times N \times P} = [(X_G W_1 + b_1) \text{Sigmoid}(X_G W_1 + b_1)] \cdot W_2 + b_2 \quad (2)$$

其中， $W$  为线性层的权重， $b$  为线性层的偏置。

- Fold：

Fold 操作就是 Unfold 的逆向操作，将提取过局部和全局信息的张量  $X \in R^{d \times N \times P}$  按照位置关系变回  $X \in R^{d \times H \times W}$ 。

## 2.4. 分类部分

为了减少模型的参数量和计算量，将 Layer5 的输入  $X \in R^{320 \times 6 \times 6}$  进行自适应平均池化(AvgPool)，之后展平为 320 个特征值，最后经过线性层和 SoftMax 之后得到 4 个模型的最终预测概率值，如公式(3)所示[13]。

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_{i=0}^3 e^{x_i}} \quad (i = 0, 1, 2, 3) \quad (3)$$

其中  $x_i$  是输入的类别所对应的特征值， $\text{Softmax}(x_i)$  为四个概率值中最大的概率值，所对应的类别即为模型的预测类别。

## 3. 实验

### 3.1. 数据收集和整理

本项研究采纳的数据集源自 Kaggle 平台，涵盖了 Eye Disease Dataset By Hritwij 和 Eye Disease Dataset 以及其他相关眼部疾病数据集的综合。为了提升模型的泛化能力和鲁棒性，采用运用随机旋转、翻转、

裁剪等多种数据增强技术来扩充和优化数据集[14]。该数据集共包含四种类型的眼部特征，其中包括三种常见的眼部疾病特征以及一种正常的眼部特征。各类特征均包含 1000 张图片，所有图像均被统一处理至  $512 \times 512$  像素尺寸。数据集分为两个子集：训练集和测试集。其中，训练集包含 3600 张图像，占全部图像的 90%，主要用于构建和训练人工智能模型。测试集则包含 400 张图像，占总数的 10%，用于评估所建立 AI 模型的性能表现[15]。这种划分方式有利于客观地评价模型在实际应用中的效果，确保研究成果的可靠性和实用性。

### 3.2. 实验环境

本研究的实验设施主要分为两大部分：计算机和智能手机。所有的模型训练，包括 MobileEDT 模型以及其他轻量级模型，均在计算机上完成，确保了训练过程的全面性与高效性。鉴于模型的最终应用目标是智能手机，因此特别重视在智能手机平台上进行模型的精确度与效率验证。所以验证这一步骤则在智能手机上实施。在进行模型训练的过程中，选用的软件环境包括 PyTorch 1.9.0 版本以及 Python 3.7.16 版本。具体的硬件配置详情请参照表 2。

**Table 2.** Hardware configuration

**表 2. 硬件配置**

| 系统           | RAM | 品牌                       | CPU  |
|--------------|-----|--------------------------|--|
| Windows 10   | 8G  | Lenovo ideapad 510-15IKB | Intel(R) Core (TM) i5-7200U CPU @ 2.50GHz (2701 MHz) |
| Android 7    | 2G  | Redmi Note 5A            | Snapdragon 425                                       |
| Android 12   | 8G  | Honor 60                 | Snapdragon 778G                                      |
| Harmony OS 3 | 8G  | HuaWei P30 Pro           | Kirin 980  |
| IOS 15.1     | 2G  | iPhone 7                 | A10  |
| IOS 16.4.1   | 6G  | iPhone 14                | A15  |

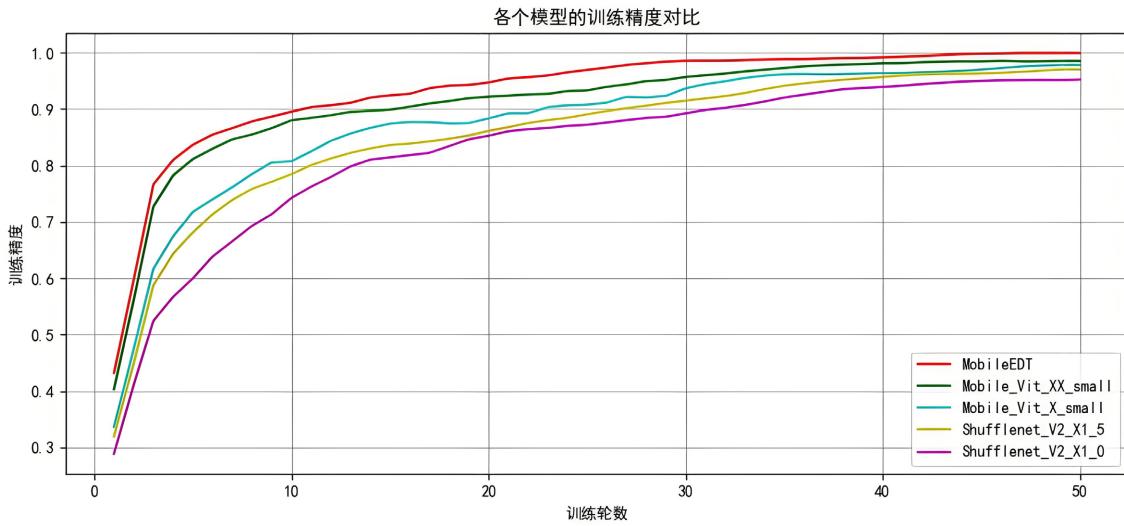
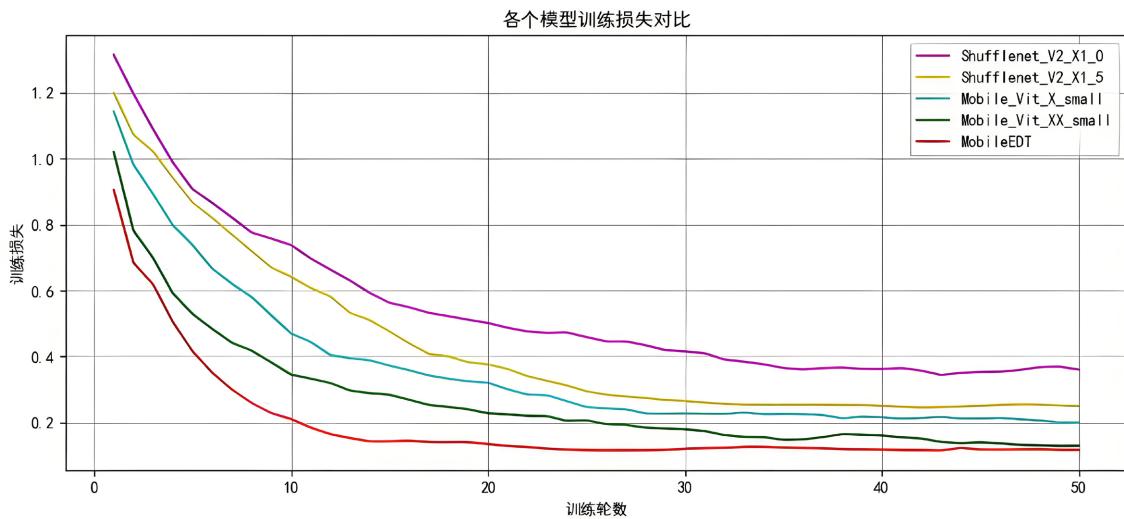
### 3.3. 实验设计与分析

#### 3.3.1. 实验设计

本研究对 MobileEDT、Mobile\_vit\_x\_small、Mobile\_vit\_xx\_small 以及 ShuffleNet\_v2\_x1\_0、ShuffleNet\_v2\_x1\_5 这五种模型的计算量(FLOPs)、参数数量和模型占用的内存进行了详细的比较[16]。此外，在智能手机上的测试环节，进一步对比了这些模型在加载时间、平均每张图像预测所需时间、预测结果的混淆矩阵以及精度指标(Kappa)等方面的表现，以全面评估各模型在实际应用环境下的效率和准确性。

#### 3.3.2. 参数设置

为确保模型性能比较的公正性，每个模型的训练过程中设置了统一的训练参数，从而避免了因参数差异而影响最终结果。训练时输入的图像分辨率固定为  $224 \times 224$  像素尺寸。训练阶段采用了基于梯度下降法的 Adam (Adaptive Moment Estimation) 优化算法进行模型优化，并设定了 0.0002 的学习率。此外，选用交叉熵(Cross-Entropy)作为损失函数[17]。整个训练过程设定为 50 个 epochs，即 50 轮完整的数据迭代，而每次迭代向模型提供 8 个样本进行训练，这是训练的 batch size 设置[18]。各模型的训练进度，包括训练精度和训练损失的演变，已在图 2 和图 3 中详细展示。

**Figure 2.** Comparison of training accuracy of each model**图 2.** 各个模型的训练精度对比**Figure 3.** Comparison of training loss for each model**图 3.** 各个模型训练损失对比

如图 2 和图 3 所示, MobileEDT 在训练期间的精度显著高于其他模型, 且其收敛速度也快于其他模型, 最终损失值低于其他模型。

### 3.3.3. 精度分析

通过混淆矩阵分析预测结果, 可以获得准确率(Accuracy)、精确率(Precision)、召回率(Recall)以及 Kappa 系数等精度评估指标[19]。混淆矩阵作为一种评估分类模型的基础工具, 通过比较预测标签与真实标签, 直观地展现了模型对于四种眼部特征的辨识力[20]。

准确率反映了分类模型预测正确结果的比例, 计算公式如公式(4)所示; 精确率则是在所有被模型判定为本类的预测中, 真正属于该类的比例, 通过计算四类眼部特征的精确率并取平均得到总体精确率, 计算方法见公式(5) [21]; 召回率是指在所有真实为本类的样本中, 模型正确预测的比例, 公式(6)给出了具体的计算公式; F1 值作为精确率和召回率的调和平均数, 其值域在 0 到 1 之间, 能够体现模型的稳定

性, F1 值越高, 表明模型的稳健性越好, 计算公式如公式(7)所示。

Kappa 系数是一个反映一致性和评估分类准确性的指标, Kappa 系数的数值越接近 1, 表示模型的分类性能越优秀。其计算公式如公式(8)所示。

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

$$F1\text{-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

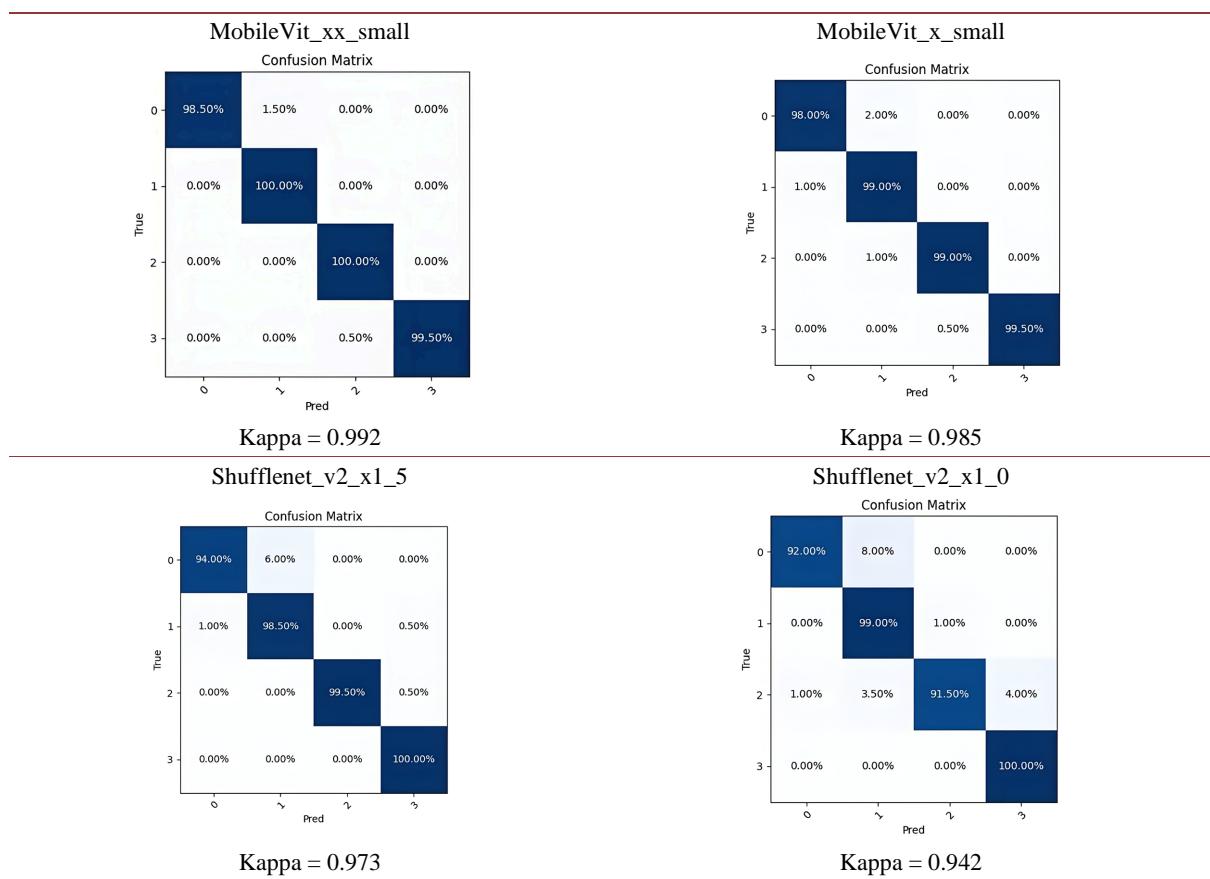
$$\text{Kappa} = \frac{\text{Accuracy} - p_e}{1 - p_e} \quad (8)$$

TP 代表实际正例, 同时被正确预测为正例的样本数; TN 指的是实际负例且被预测为负例的样本数目; FP 表示实际负例却被错误预测为正例的样本数; FN 则指实际正例却未被预测为正例的样本数。此外,  $p_e$  值是所有类别各自的“实际值与预测值乘积之和”相对于“样本总数平方”的比例。

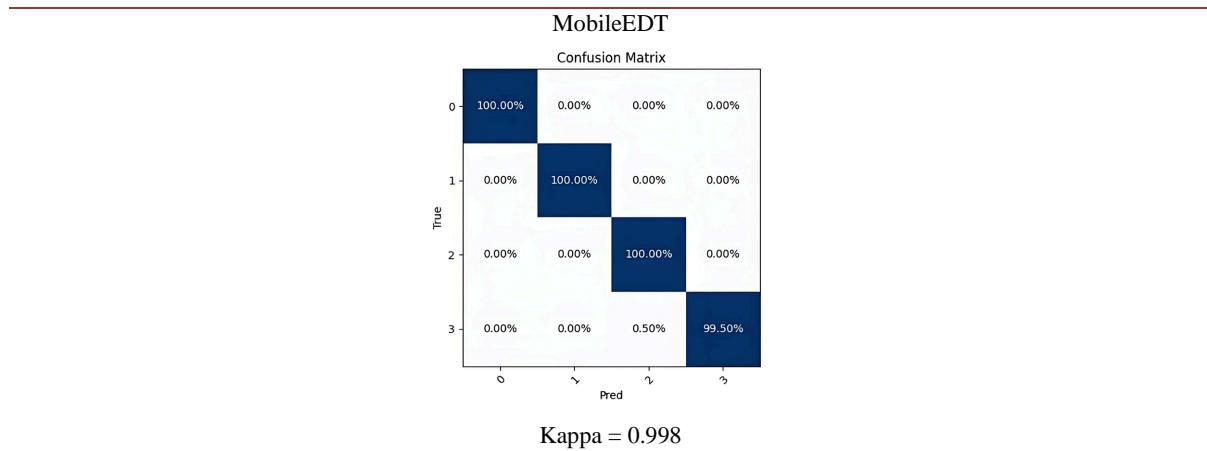
各模型的混淆矩阵及 Kappa 系数已在表 3 中列出, 而准确率(Accuracy)、精确率(Precision)、召回率(Recall)的具体数值则详见表 4。

**Table 3.** Confusion matrix and Kappa coefficients for each model

**表 3.** 各个模型的混淆矩阵和 Kappa 系数



续表

**Table 4.** Comparison of accuracy metrics for each model**表 4.** 各个模型的精度指标对比

| 模型                 | Accuracy     | Precision    | Recall       | F1-Score     |
|--------------------|--------------|--------------|--------------|--------------|
| MobileVit_xx_small | <b>0.995</b> | <b>0.995</b> | <b>0.995</b> | <b>0.995</b> |
| MobileVit_x_small  | <b>0.989</b> | <b>0.989</b> | <b>0.989</b> | <b>0.989</b> |
| Shufflenet_v2_x1_5 | <b>0.980</b> | <b>0.982</b> | <b>0.980</b> | <b>0.981</b> |
| Shufflenet_v2_x1_0 | <b>0.956</b> | <b>0.959</b> | <b>0.956</b> | <b>0.957</b> |
| MobileEDT          | <b>0.999</b> | <b>0.999</b> | <b>0.999</b> | <b>0.999</b> |

表 3 展示了 MobileEDT 与其他模型在 Kappa 系数和混淆矩阵方面的评估结果。数据显示, MobileEDT 在眼部特征分类任务中取得了最高的 Kappa 系数, 达到了 0.998, 而在测试的模型中 Shufflenet\_v2\_x1\_0 的 Kappa 系数最低, 为 0.942。另外, MobileVit 模型中 MobileVit (xx\_small) 的 Kappa 系数高于 MobileVit (x\_small), 而在 Shufflenet 系列模型中, Shufflenet\_v2\_x1\_5 的 Kappa 系数高于 Shufflenet\_v2\_x1\_0。表 4 进一步证实了这五种模型的精度都非常高, 其中 MobileEDT 的精度最高, 达到了 0.999。综合上述结果, 可以得出结论, MobileEDT 展现出了极高的性能精度。

### 3.3.4. 效率分析

已训练模型被部署至不同型号的智能手机上, 以评估其模型效率。评估指标包括模型的计算量(FLOPs)、参数数量、模型所占的内存空间、每张图像的识别时间(ART), 以及智能手机加载模型至内存所需的时间(LT), 如公式(9)所示[22]。

$$\text{ART} = \frac{\text{TRT}}{\text{Amount}} \quad (9)$$

其中 TRT 代表识别整个测试集图像所需的总时间, 而 Amount 则是测试集中图像的总数。实验结果详见表 5 和表 6, TT 表示加载模型时间(LT)与平均每张图像检测时间(ART)的和。

从表 5 和表 6 的整体数据来看, MobileEDT 展现了最优的加载模型时间(LT), 在 iPhone 14 上仅需 50 毫秒, 同时拥有最小的参数量(1.27M)和最小的内存占用(3.72MB)。此外, MobileEDT 的计算量(FLOPs)在测试的型号中位列第二低, 仅高于 ShuffleNet\_v2\_x1\_0 模型。

**Table 5.** Comparison of the overall efficiency of the models  
**表 5.** 各个模型的总体效率对比

| 模型                   | 品牌(RAM)             | LT (ms) | ART (ms) | TT (ms) |
|----------------------|---------------------|---------|----------|---------|
| MobileVit (xx_small) | Redmi Note 5A (2G)  | 478     | 283      | 761     |
|                      | HuaWei P30 Pro (8G) | 92      | 52       | 144     |
|                      | HONOR 60 (8G)       | 102     | 62       | 164     |
|                      | iPhone 7 (2G)       | 121     | 112      | 233     |
|                      | iPhone 14 (6G)      | 105     | 55       | 160     |
|                      | Redmi Note 5A (2G)  | 550     | 144      | 694     |
| MobileVit (x_small)  | HuaWei P30 Pro (8G) | 116     | 37       | 153     |
|                      | HONOR 60 (8G)       | 128     | 41       | 169     |
|                      | iPhone 7 (2G)       | 93      | 57       | 150     |
|                      | iPhone 14 (6G)      | 83      | 50       | 133     |
|                      | Redmi Note 5A (2G)  | 643     | 107      | 750     |
|                      | HuaWei P30 Pro (8G) | 146     | 30       | 176     |
| Shufflenet_v2_x1_5   | HONOR 60 (8G)       | 160     | 32       | 192     |
|                      | iPhone 7 (2G)       | 124     | 99       | 223     |
|                      | iPhone 14 (6G)      | 115     | 52       | 167     |
|                      | Redmi Note 5A (2G)  | 732     | 572      | 1304    |
|                      | HuaWei P30 Pro (8G) | 168     | 100      | 268     |
|                      | HONOR 60 (8G)       | 179     | 118      | 297     |
| Shufflenet_v2_x1_0   | iPhone 7 (2G)       | 153     | 174      | 327     |
|                      | iPhone 14 (6G)      | 125     | 73       | 198     |
|                      | Redmi Note 5A (2G)  | 476     | 117      | 593     |
|                      | HuaWei P30 Pro (8G) | 90      | 35       | 125     |
|                      | HONOR 60 (8G)       | 101     | 40       | 141     |
|                      | iPhone 7 (2G)       | 57      | 55       | 112     |
| MobileEDT            | iPhone 14 (6G)      | 50      | 29       | 79      |

**Table 6.** Calculations, number of parameters and memory size of the model for each model  
**表 6.** 各个模型的计算量，参数量以及模型所占内存大小

| 模型                  | 计算量(FLOPs) | 参数量    | 内存大小    |
|---------------------|------------|--------|---------|
| MobileVit(xx_small) | 261.6 M    | 1.28 M | 3.76 MB |
| MobileVit(x_small)  | 720.17 M   | 2.23 M | 7.52 MB |
| ShuffleNet_v2_x1_5  | 306.05 M   | 3.5 M  | 9.68 MB |
| ShuffleNet_v2_x1_0  | 152.39 M   | 2.28 M | 4.97 MB |
| MobileEDT           | 231.66 M   | 1.27 M | 3.72 MB |

对于不同型号的智能手机，运行内存的容量直接决定了模型在 LT 和 ART 上的性能。拥有更大运行内存的智能手机通常能以更短的时间完成任务。**表 5** 显示，运行内存较小的 Redmi Note 5A (2G) 在 LT 和 ART 上耗时最长，而即便是相同运行内存的手机，由于 CPU 的差异，其 LT 和 ART 也会有所区别，例如搭载 Kirin 980 的 Huawei P30 Pro (8G) 比配备 Snapdragon 778G 的 Honor 60 (8G) 所需时间略少。再者，在 LT 和 ART 的比较中，iPhone 智能手机相较于 Android 智能手机表现出更快的处理速度。总体而言，尽管不同型号的智能手机在 LT 和 ART 上存在差异，但它们的总体运行时间(TT)均控制在 2 秒以内，这在实际应用中是可以接受的。

综合以上对比分析，MobileEDT 在各类智能手机上的整体性能表现超越了其他轻量级模型。

## 4. 结论与展望

本研究通过整合数据，设计了一种名为 MobileEDT 的新型智能手机分类模型，旨在应对当前眼部疾病诊断的挑战。该模型能够迅速检测眼部疾病，为患者提供及时的医疗援助。MobileEDT 采用卷积神经网络提取局部特征，随后利用 Transformer 捕获全局特征，并通过卷积层将局部与全局特征融合。最终，特征张量被送入全连接层进行预测。

综上所述，本研究提出了一种创新的眼部疾病诊断模型 MobileEDT，不仅展现出高性能、高精度和低参数量等优点，还展示了其在移动健康领域的巨大潜力，为未来的医疗技术发展和应用提供了有力支持。

尽管本研究成功实现了对眼部特征的识别，但由于实际眼部疾病的数据集有限，目前仅涵盖了四种眼部特征。未来，通过收集更广泛、更多样化的眼部疾病数据，有望进一步扩展和优化 MobileEDT 模型，提升其诊断能力和适用范围，为眼科医疗领域带来更大的贡献。

## 参考文献

- [1] 李朝林. 基于移动设备的眼底疾病智能检测算法及应用研究[D]: [硕士学位论文]. 贵阳: 贵州大学, 2023.
- [2] 黄林哲, 刘力学, 吴雨璇, 等. 基于深度学习和智能手机的眼病预防与远程诊疗[J]. 眼科学报, 2022, 37(3): 230-237.
- [3] 贺晨. 基于深度学习的白内障病变图像分级方法研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2020.
- [4] 张婉芸. 基于人工智能与羟甲基化修饰的葡萄膜炎诊断及发病机制研究[D]: [硕士学位论文]. 重庆: 重庆医科大学, 2023.
- [5] 郑春荣. 让眼睛变大、突出的甲状腺相关眼病[J]. 食品与健康, 2024, 36(1): 38-39.
- [6] 宫阿娟. 基于 ResNet 深度神经网络构建眼部疾病分类诊断模型的研究[J]. 医药论坛杂志, 2024, 45(4): 379-383.
- [7] Tivive, F.H.C. and Bouzerdoum, A. (2006) Rotation Invariant Face Detection Using Convolutional Neural Networks. *Neural Information Processing*, Hong Kong, 3-6 October 2006, 260-269. [https://doi.org/10.1007/11893257\\_29](https://doi.org/10.1007/11893257_29)
- [8] 于静. 基于苹果移动终端设备的眼疾诊断方法的研究[D]: [硕士学位论文]. 兰州: 兰州大学, 2012.
- [9] Prasad, K., Sajith, P.S., Neema, M., Madhu, L. and Priya, P.N. (2019). Multiple Eye Disease Detection Using Deep Neural Network. *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, Kochi, 17-20 October 2019, 2148-2153. <https://doi.org/10.1109/tencon.2019.8929666>
- [10] Rahman, M.M., Islam, M.S., Ara Jannat, M.K., Rahman, M.H., Arifuzzaman, M., Sassi, R., et al. (2020). EyeNet: An Improved Eye States Classification System Using Convolutional Neural Network. *2020 22nd International Conference on Advanced Communication Technology (ICACT)*, Phoenix Park, 16-19 February 2020, 84-90. <https://doi.org/10.23919/icact48636.2020.9061472>
- [11] Mehta, S. and Rastegari, M. (2021) MobileViT: Light-Weight, General-Purpose, and Mobile-Friendly Vision Transformer. arXiv: 2110.02178. <https://doi.org/10.48550/arXiv.2110.02178>
- [12] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.

- 
- [13] 徐金涛. 基于深度学习的眼底多疾病分类研究[D]: [硕士学位论文]. 上海: 东华大学, 2024.
  - [14] Glaret Subin, P. and Muthukannan, P. (2022) Optimized Convolution Neural Network Based Multiple Eye Disease Detection. *Computers in Biology and Medicine*, **146**, Article 105648.  
<https://doi.org/10.1016/j.combiomed.2022.105648>
  - [15] Zhang, X., Hu, Y., Xiao, Z., Fang, J., Higashita, R. and Liu, J. (2022) Machine Learning for Cataract Classification/Grading on Ophthalmic Imaging Modalities: A Survey. *Machine Intelligence Research*, **19**, 184-208.  
<https://doi.org/10.1007/s11633-022-1329-0>
  - [16] Ma, N., Zhang, X., Zheng, H. and Sun, J. (2018) ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. *Computer Vision—ECCV 2018*, Munich, 8-14 September 2018, 122-138.  
[https://doi.org/10.1007/978-3-030-01264-9\\_8](https://doi.org/10.1007/978-3-030-01264-9_8)
  - [17] 娄茹珍, 徐丽, 蒋正乾, 等. 基于卷积神经网络的眼疾识别算法[J]. 无线电工程, 2021, 51(11): 1202-1207.
  - [18] Rao, P.S. and Sreehari, S. (2012) Neural Network Approach for Eye Detection. *Computer Science & Information Technology*, **5**, 269-281. <https://doi.org/10.5121/csit.2012.2228>
  - [19] Ahmed, I.O., Ibraheem, B.A. and Mustafa, Z.A. (2018) Detection of Eye Melanoma Using Artificial Neural Network. *Journal of Clinical Engineering*, **43**, 22-28. <https://doi.org/10.1097/jce.0000000000000250>
  - [20] 刘云. 基于深度学习的视网膜 OCT 图像分层与疾病筛查研究[D]: [硕士学位论文]. 济南: 山东大学, 2018.
  - [21] 任章军, 余进海, 桑泽曦, 等. 人工智能深度学习在眼眶病及眼肿瘤疾病诊疗中的应用研究现状[J]. 眼科新进展, 2024, 44(2): 163-168.
  - [22] 张晓青, 刘小舟, 陈登. 面向移动端图像分类的轻量级 CNN 优化[J]. 计算机工程与设计, 2024, 45(2): 436-442.