

基于神经网络的手势识别技术在直升机上的应用研究

谷小乐¹, 余强胜², 陈玲¹

(1. 中国直升机设计研究所, 江西 景德镇 333001; 2. 空装驻景德镇地区军事代表室, 江西 景德镇 333001)

摘要: 随着直升机作战需求的不断提升, 采用传统的按键方式作为人机交互形式来操作大屏显示器越来越不能适应需求。将触控手势识别技术应用于直升机座舱人机交互界面领域具有一定的优势, 而传统的触控手势识别基本都是在固定的触控板上进行触控操作, 场景受限严重, 并不能很好地适用于直升机机载环境。基于此, 本文设计了一种一体式大触控显示器框架, 提出了直升机机载环境下基于神经网络的全屏幕触控手势识别的技术方案, 其中神经网络可以自动提取适合分类的手势特征, 采集多人的手势书写习惯进行样本的学习训练。通过对实际产品的测试发现, 识别效果基本满足设计需求。

关键词: 手势识别; 神经网络; 直升机

中图分类号: TP391.4

文献标识码: A

DOI: 10.12230/j.issn.2095-6657.2023.11.037

触控显示器与传统设备的最大不同在于, 它在作为信息显示设备的同时, 也可作为控制设备/输入设备, 即实现了输出和输入设备的集成和统一。使用显示区域和控制区域相结合的触控屏设备, 飞行员则可以直接在同一区域内完成信息的读取和飞行操作任务^[1]。基于此, 本文提出了一种适用于直升机座舱的一体式大触控显示器框架, 并针对这一显示器框架设计了基于神经网络的全屏幕触控手势识别的技术。

1 触控屏技术在直升机上应用的优势

1.1 提高手、眼操作的协调性

引入触控屏的驾驶舱, 可以使飞行员在执行飞行任务时的注意力集中于同一区域。不仅包括飞行员眼睛的注视点, 还包括飞行员用手指在触控屏上进行操作。因操控注意力和视觉注意力均集中于同一区域, 飞行员在操控过程中可以在同一位置获得信息的反馈, 而无需移动视觉注意力, 因而该操作方式降低了对飞行员手、眼协调工作能力的要求, 使得飞行员可以更加容易地执行任务。

1.2 降低脑力负荷

触控屏技术实现了显示与控制区域的结合, 为实现飞行员获得对操控动作的即时性信息反馈提供了可能。操作传统设备时, 飞行员完成某个操控动作后, 需要将视线移动到该控制器所对应的显示器上进行相关信息确认, 该过程会占用部分脑力负荷。采用触控屏设备, 飞行员在完成输入的同时, 在同地即可获得信息的反馈, 省去判别信息显示和控制操作相对应的脑力消耗。

1.3 缩短操作时间

触控屏技术可大量缩短操作时间。因显示和控制相集成于同一器件, 飞行员的注视焦点和操作对象位于同一区域, 免去了采用传统显示和操控设备的驾驶舱所需的视线焦点在控制区域与显示区域上来回移动的情况。

早期的触控手势识别需要额外的触控设备, 成本较高, 交互形式不自然, 且识别方法侧重于图像识别的角度, 将触控手势转化为图形, 利用相关算法进行图形特征提取, 并对图形进行分类。该类方法的灵活性较低, 且存在一定的局限性, 通常将触控手势转化成图像需要耗费一定的算力与时间, 增加了硬件的负担, 同时也增加了手势识别的时间^[2]。

随着大数据研究的兴起, 机器学习变得异常火热, 人工神经网络(Artificial Neural Network)作为其中代表, 其特点是模拟人脑的信息传递机制。神经网络分为生物神经网络以及人工神经网络^[3]。神经网络技术已经广泛地应用于图像处理、语音识别等领域, 本次研究采用的神经网络算法, 是将屏幕触控手势的坐标经过归一化后直接作为神经网络的输入, 输出即为识别的结果, 训练完成后的网络模型可以适用于多种尺寸的屏幕, 在更换显示平台后不需要进行再次的适配, 提高了触控手势识别的通用性和智能性, 既立足于最新的理论, 又具有实际的应用背景。对于嵌入式平台, 杨晨奕等人提出过轻量化神经网络手势识别方法, 但对于面向机载环境下的显控系统, 并没有有效的应用^[4]。

2 一体式大触控显示器框架

一体式大触控显示器的原理框图:

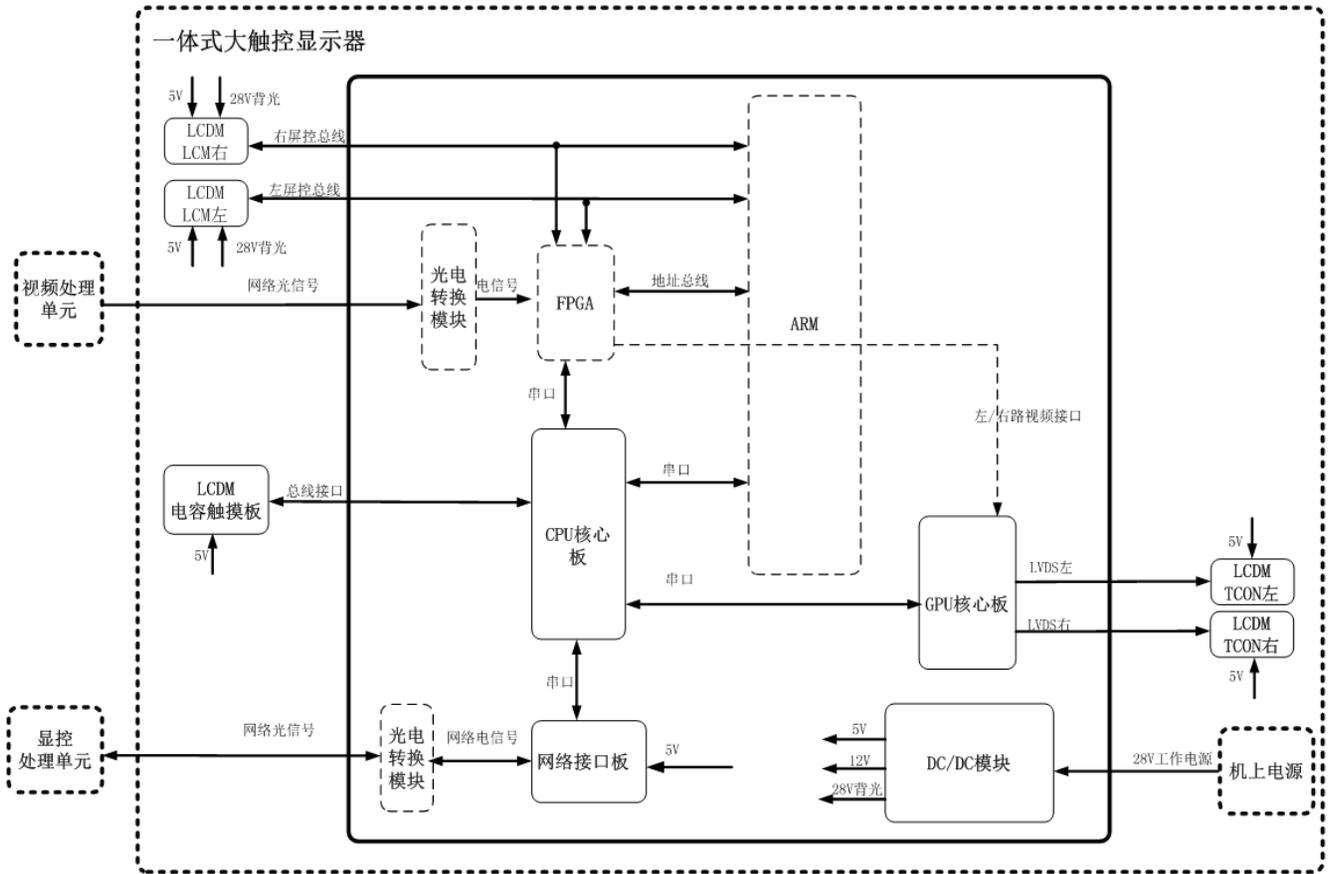


图1 一体式大触控显示器的原理框图

一体式大触控显示器主要划分为以下功能模块：底板、CPU 核心板、GPU 核心板、网络接口板（FCM）、液晶显示模块（LCDM）和电源转换模块。

底板的核心架构采用 ARM+FPGA 的形式，主要实现一路视频信号的光电转换及视频解码，实现双冗余网络信号的光电转换,实现数据总线、左右屏控数据总线通信信号的转发。同时，底板实现各个不同模块之间的信号传递。

CPU 核心板作为显示器的通讯控制中心，实现数据总线、左右屏控数据总线、触控数据总线、网络等通信信号的解析控制，实现控制 GPU 作图。

GPU 核心板作为显示器的视频画面管理中心，根据 POP 要求自主产生视频画面，同时可实现与外视频信号叠加，以左右 LVDS 格式的视频信号送给液晶显示模块进行显示。

网络接口板（FCM）主要实现双冗余网络信号的收发，通过串口与 CPU 核心板实现通信。

液晶显示模块（LCDM）由 LCM 左、LCM 右、TCON 左、TCON 右和电容触控板组成。液晶显示模块通过 TCON 板接收 GPU 核心板 LVDS 信号实现画面的显示，同时通过触控数据总线实现触控信号的控制，通过屏控数据总线实现亮度、日夜模式切换等控制。

电源转换模块主要实现机上工作电源 28V 电压的转换，转换成各个模块所需要的电压，包括 5V、12V 和背光 28V。

3 手势数据与神经网络模型建立

3.1 手势数据集建立

在建立触控手势识别神经网络模型之前，需要分析触控手势数据，明确触控手势数据的特征数，以及手势的类型数。因此公开可用的数据集对研究人员非常重要，尤其是机载场景下的手势数据集更为重要，但是并没有公开权威的机载环境下触控手势数据集。为了便于课题研究的进行，本次研究开展了数据库建立工作，本文中的数据均为仿真环境下，通过大屏触控显示器采集所得，真实可靠。预处理操作中主要介绍了手势关键点的获取。

本数据集中的数据包含了常见的 6 种手势,分别为 O、M、V、S、N 和 W。受试人员为某试飞大队直升机飞行员。受试人数为 10。手势采集系统通过 C 语言开发，采集设备为 ViewSonic 触控显示器、触控计算机。采集步骤如下：首先在触控计算机上打开触控采集程序，设置采样率，每人每种手势分别书写 10 次。

每组手势数据均采样 20 个点,每个点包含 x 与 y 两个坐标，

所以每个样本包含 40 个特征，因此 BP 神经网络的输入节点数为 40。输出则对应了手势的类型，共有 6 种手势，因此输出节点数为 6。

在获取了所有触控数据之后，根据屏幕尺寸，将触控数据进行归一化，归一化能够将不同数量级的数据压缩到同一量纲，且提升模型的收敛速度和模型精度，常见的归一化方法有 min-max 标准化和 Z-score 标准化两种。

min-max 标准化，也称为离差标准化，是对原始数据的线性变化，使得结果映射到 [0, 1] 之间。转换函数如下所示：

$$x^* = \frac{x - \min}{\max - \min} \quad (1)$$

其中 max 为样本数据的最大值，min 为样本数据的最小值。这种方法有个缺陷，就是当有新的数据加入时，可能导致 max 和 min 变化，需要重新定义。

本次研究由于数据的上下界均为已知数，所以采用了 min-max 标准化对数据进行预处理。

3.2 BP 神经网络的手势识别模型建立

神经网络可以自己学习，具备自动组织的能力，同时还具备抗噪能力，对不完整的模式也可以很好地解决，其中广泛应用于 BP (back propagation) 神经网络是前馈型的神经网络，所采用的传播算法为逆向传播，根据输入值和期望值之间的误差进行权值的更新。神经网络算法具有并行分布处理、非线性映射、所需特征参数简单明确的特性，具有较高的识别准确率。触控手势识别中，采用 BP 神经网络算法进行复杂手势识别工作。

BP 神经网络是一种多层前向型神经网络，可学习和存储大量输入输出模式映射关系，而无需事前揭示描述这种映射关系的数学方程，这一特点非常适合手势轨迹的计算。

以归一化特征向量作为输入，对应的手势作为输出，网络隐含层采用双隐层，隐含层神经元转移函数取双曲正切 S 型传递函数 tansig，输出层神经元转移函数取线性传递函数 purelin，学习规则为 trainlm。

常见的 BP 神经网络包含输入层、隐含层以及输出层，隐含层超过 1 的神经网络又被称为深度神经网络，输入层以及输出层节点数需要根据数据集来设置，隐含层层数以及节点数则需要根据网络的训练结果来人为调整。

本次研究通过调整隐含层层数、隐含层神经元个数，以及学习率来优化网络的识别精度。

通过神经网络模型完成了对触控数据集的训练与测试，将数据集划分为训练集、验证集和测试集，分别用于模型的训练、优化与测试。通过训练过程的迭代图，可以发现数据集能够较好地反映触控手势特征。

误差柱状图反映了数据集整体的误差分布，可以发现大部分的误差均集中在 0 位附近，由此可见模型的泛化能力较强，不会出现个别误差过大的情况。ROC 曲线体现了曲线上各点对同一信号刺激的反应。AUC 表示 ROC 曲线下的面积，主要用于衡量模型的泛化能力，即分类器的性能，AUC 等于 1 是完美分类器，采用这个预测模型时，存在至少一个阈值能得出完美的预测。越接近 1 越优秀。

通过查看手势识别模型的混合图可以发现模型的训练精度超过 95%，在验证过程中精度也在 95% 以上。

4 机载大屏基于神经网络的触控手势识别

在完成触控手势识别网络的训练后，即可将手势识别模型移植到产品中。本次研究中，采用了国产化硬件平台，显示程序与触控识别程序通过 DD 文件进行交互，显示程序将触控数

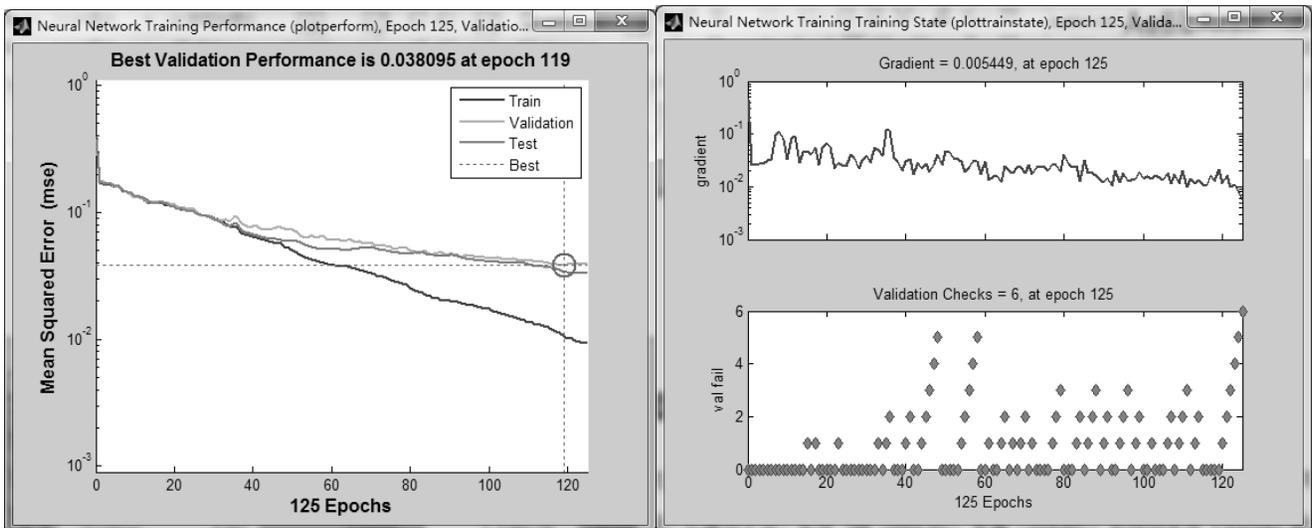


图 2 模型迭代过程

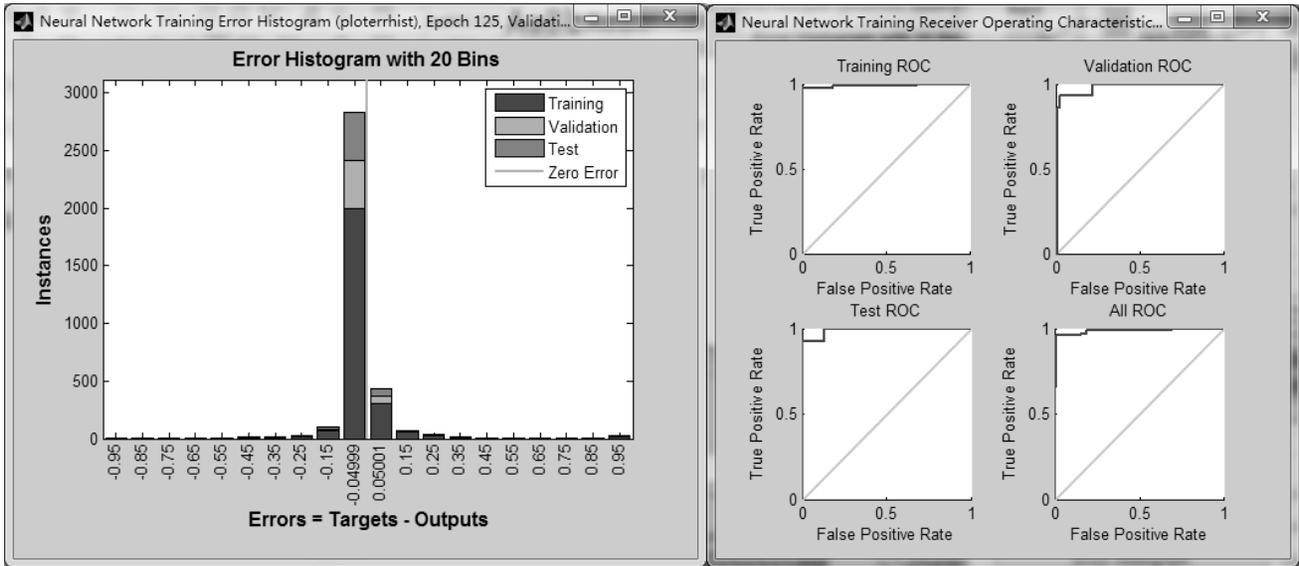


图3 误差分布及 ROC 曲线

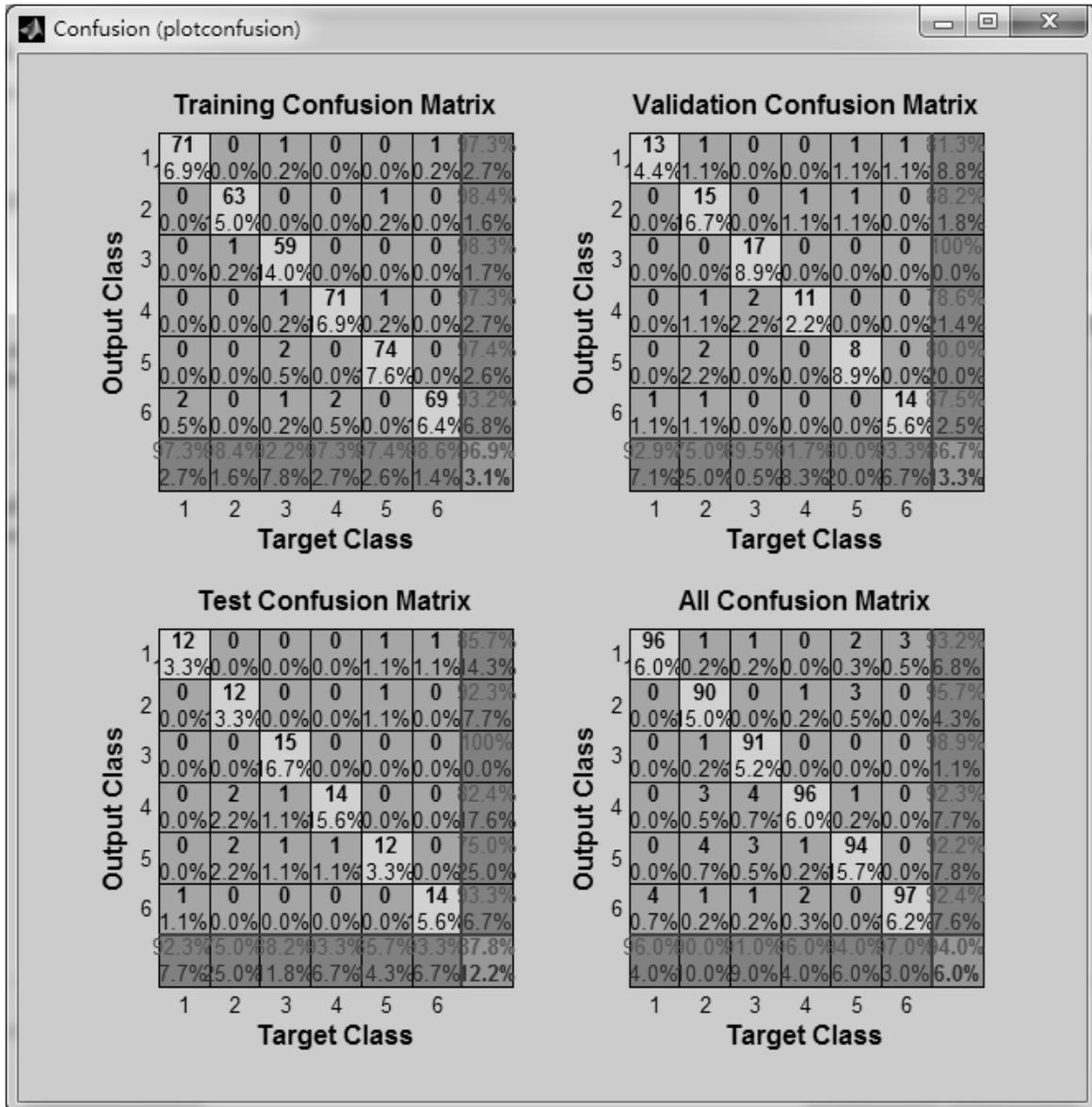


图4 误差混合图

据存入 DD 文件中，触控识别程序通过访问 DD 文件进行手势识别，并将识别结果写入 DD 文件中，画面显示程序接收识别结果。

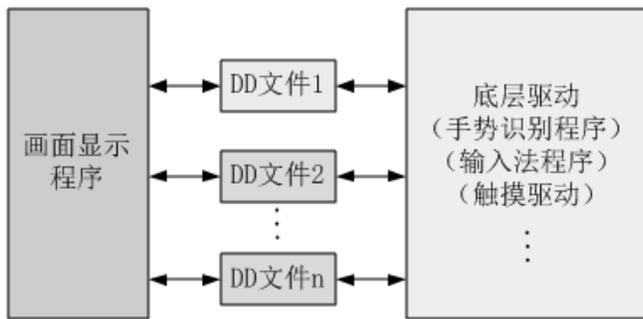


图 5 触控识别框架图

为测试触控手势的识别情况，需要编写对应的显示软件，绘制对应的触控区域，实时打印触控坐标，并打印识别结果。配置对应的 DD 文件，将触控数据按数据集要求打包通过 SendBuf 写入 DD 文件。识别结果通过判断 RecBuf 中的数据显示不同的结果。1 到 6 分别代表 6 种不同的手势，根据结果，画面显示程序可以做出不同的响应，切换画面或触发某些按钮。



图 6 触控显示软件界面

手势识别模型通过神经网络算法模型实现，将其转换成 C++ 代码移植产品底层驱动程序中。其中输入层包含 40 个节点，输出为 6 种手势的识别概率，可以通过判断概率大小得到最终的结果。

本课题所开发的机载触控手势识别系统操作简单、使用便

捷，可以实现机载场景下相应的触控手势识别，并且采集数据模块也便于使用，但实时在线预测手势时存在预测错误的情况，可能是由于采集数据时卡顿等原因，下一步工作中会保证在此基础上进行改进，以达到一个满意的性能。

5 结论

综上所述，手势识别作为一种简单、自然的人机交互方式，一直备受研究人员的关注并对其进行各种算法的设计。鉴于此，本文提出了一种效果较好的基于神经网络的手势识别方法，并搭建了机载场景下基于神经网络的手势识别系统。本次研究的主要工作总结如下：(1) 采集了机载环境下大型触控屏 6 种常见的手势，并对数据集进行了相应的预处理。(2) 开展了基于神经网络的触控手势识别研究，并根据识别结果对网络的结构进行了优化，使得识别效果满足需求。(3) 开发了可适用于国产化硬件平台的手势识别程序，操作简单，具有良好的人机交互性能，为下一步所要开展的优化算法及数据质量评估提供了基础平台。

参考文献：

- [1] 陆虎敏. 飞机座舱显示与控制技术 [M]. 北京: 航空工业出版社, 2015.
- [2] 何小磊. 基于深度信息和彩色图像的手势识别 [D]. 武汉: 华中科技大学, 2012.
- [3] 朱大奇, 史慧. 神经网络原理及应用 [M]. 北京: 科学出版社, 2006.
- [4] 杨晨奕, 何玉青, 赵俊媛, 等. 面向嵌入式平台的轻量化神经网络手势识别方法 [J]. 强激光与粒子束, 2022, 34 (03): 138-146.

作者简介：谷小乐（1991-），男，江西抚州人，工程师，硕士研究生，主要从事直升机航电系统设计研究。