

文章编号: 1672-8785(2023)12-0032-09

基于卷积神经网络的胸环靶面畸变校正算法

姜宗林¹ 陈小林¹ 李鹏飞¹ 王东鹤¹ 吴志佳¹ 王 虹²

(1. 中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国联合网络通信有限公司长春市分公司, 吉林 长春 130061)

摘 要: 实弹射击是部队的基础军事训练项目。现有报靶系统中基于计算机视觉的弹孔识别定位系统由于具有快速、精确、安全、人员成本低等优点而被广泛应用到该项目中。然而, 计算机视觉系统处理的图像通常受镜头加工工艺以及相机轴向与被测对象所在平面不垂直的影响, 导致被测对象的图像产生畸变, 最终会给弹孔坐标位置的精准定位带来误差。为了提高基于计算机视觉的自动报靶系统的报靶精度, 提出一种基于卷积神经网络的畸变校正算法, 只需一张胸环靶面的模板图像即可模拟出大量训练数据集。训练完成后, 输入一张畸变图片就可以得到该图片的畸变参数, 并利用该参数完成对图像的畸变校正。与传统校正算法的对比结果表明, 该算法校正效果较好, 有利于提升基于计算机视觉的自动报靶系统的报靶精度。

关键词: 机器视觉系统; 图像畸变; 校正算法; 卷积神经网络

中图分类号: TP391 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2023.12.005

Chest Ring Target Distortion Correction Algorithm Based on Convolutional Neural Network

JIANG Zong-lin¹, CHEN Xiao-lin¹, LI Peng-fei¹, WANG Dong-he¹, WU Zhi-jia¹, WANG Hong²

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;

2. Changchun Branch of China United Network Communication Co., LTD., Changchun 130061, China)

Abstract: Live firing is a basic military training item of the army. The bullet hole identification and positioning system based on computer vision in the existing target detection system is widely used in this project because it is fast, accurate, safe and it has a low personnel cost. However, the image processed by the computer vision system is usually affected by the lens processing technology. In addition, the axis of the camera is not perpendicular to the plane where the measured object is located, so the image distortion of the measured object is resulted in, and the accurate positioning error of the coordinate position of the bullet hole is generated. In order to improve the accuracy of automatic target reporting system based on computer vision, a distortion correction algorithm based on convolutional neural network, which can simulate a large number of training data

收稿日期: 2023-06-29

作者简介: 姜宗林(1996-), 男, 吉林长春人, 硕士, 主要研究方向为智能图像处理与识别、目标检测与跟踪技术等。E-mail: 280657858@qq.com

sets with only one template image of the chest circle target surface is proposed in this paper. When the training is completed, the distortion parameters of the image can be obtained by inputting a distorted image, and the image distortion correction can be completed by using the parameters. Compared with the traditional correction algorithm, the results show that the algorithm has a good correction effect and is conducive to improving the accuracy of the automatic target detection system based on computer vision.

Key words: machine vision system; image distortion; correction algorithm; convolutional neural network

0 引言

实弹射击是部队开展的基础训练项目。在传统的射击训练中,一般需要专门配备报靶人员进行人工报靶。这种方法效率低下且不安全,正逐渐被各类自动报靶系统取代。在众多不同类型的自动报靶系统中,基于计算机视觉技术的弹孔识别及定位系统凭借成本低、使用方便的特点被广泛应用^[1-2]。但是,基于计算机视觉技术的报靶系统在具体投入使用时仍然面临着需要克服的阻碍。计算机视觉系统的图像通常被镜头加工工艺以及相机轴向与被测物体所在平面的不垂直影响^[3-4],导致系统用来处理分析的靶面图像产生畸变,最终给测量结果带来误差。此外,计算机视觉系统的图像畸变问题无法在系统设计时提前避免,对弹孔定位的准确性产生了极大的影响。现有自动报靶系统的靶面校正方法是选取靶面的四个点进行透射变换。但是在室外射击环境中,由于靶面一直晃动,所以会产生很大误差。另一种方法是通过检测靶纸的四个顶点来进行自动校正,但是场景局限性较大,只能适用于某一种靶纸;部分靶板没有靶纸的白色边框或者靶纸破损便无法校正^[5-7]。因此,本文针对胸环靶面的畸变图像提出了一种基于卷积神经网络的校正算法。实验结果表明,该算法校正效果较好,有利于提升报靶精度。

1 计算机视觉系统的畸变模型

计算机视觉系统的线性成像模型如图 1 所示。其中,三维空间点 M 在图像平面上的投影成像点 $m_i(x_i, y_i)$ 表示 M 在像平面坐标系下的实际投影点的坐标, $m_r(x_r, y_r)$ 表示 M 在像平面坐标系下的理想投影点的坐标。要建立图

像非线性畸变的数学模型,就需理解计算机视觉系统成像模型中世界、相机、图像和像素四个平面坐标系之间的关系。

世界坐标系是整个成像系统的基准坐标系。将该坐标系作为基准,对置于这个空间的所有物体包括相机进行位置信息的描述,坐标值用 (X_w, Y_w, Z_w) 表示。相机坐标系的原点 O 在光心位置,即小孔成像的中心。 X 轴和 Y 轴平行于图像坐标系的 X 轴和 Y 轴, Z 轴平行于相机光轴,其坐标值用 (X_c, Y_c, Z_c) 表示。图像坐标系是二维平面,其中心为坐标原点 O_i , X 轴和 Y 轴与平面的两条垂直边平行,其坐标值用 (x, y) 表示,坐标系中的物理单位为 mm。像素坐标系与图像坐标系处在同一平面上,其原点 O_0 位于图像的左上角,其坐标值用 (u, v) 表示。计算机视觉系统采集的图像经过模数转换,每幅图像的像素值存储在 $M \times N$ 的数组之中^[10]。

在像素坐标系中,无法反映像素在图像位置中的物理尺寸。因此,在图像非线性畸变模型中用 (x, y) 表示图像像素点的物理坐标。畸变校正也是在图像坐标系上完成的。图 2 反映了两个坐标系之间的关系。

像素、图像坐标系之间的转换关系如下:

$$\begin{cases} u = \frac{x}{dx} + u_0 \\ v = \frac{y}{dy} + v_0 \end{cases} \quad (1)$$

用齐次坐标将式(1)表示成矩阵形式:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

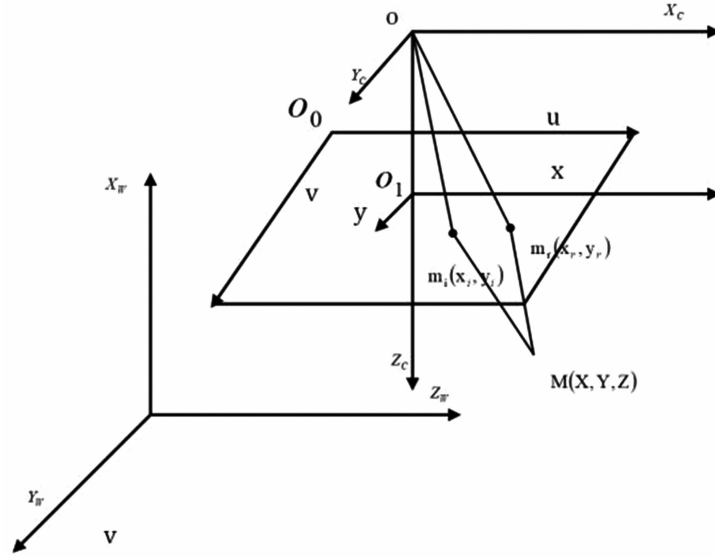


图 1 实际的机器视觉成像系统成像模型

式中, 像素坐标系中的坐标 (u_0, v_0) 是图像坐标系的原点坐标, dx 和 dy 分别为图像坐标系中每个像素在两坐标轴上的物理尺寸。

相机、图像坐标系之间的转换关系如下:

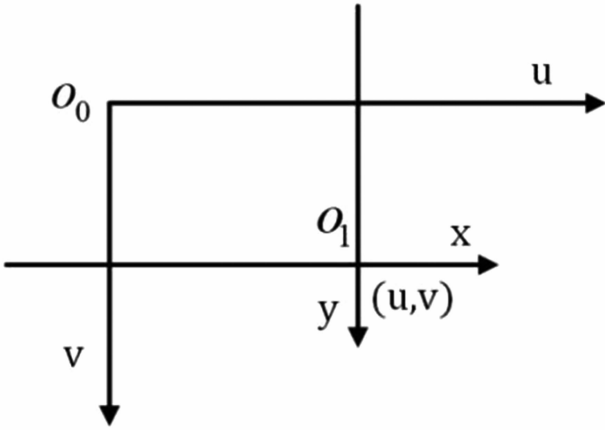


图 2 像素、图像坐标系之间的关系图

$$\begin{cases} x = \frac{fX_c}{Z_c} \\ y = \frac{fY_c}{Z_c} \end{cases} \quad (3)$$

用齐次坐标将上述关系表示成矩阵:

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (4)$$

相机坐标系与世界坐标系的变换关系如下:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (5)$$

式中, R 为旋转矩阵, 由三个轴的 3×3 旋转矩阵相乘得到; t 为三维平移向量 (t_x, t_y, t_z) 。

将式(1)~式(5)综合起来表示为

$$\begin{aligned} Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \\ &= KM_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (6) \end{aligned}$$

式中, a_x 和 a_y 分别是图像坐标系 X 方向和 Y 方向上的尺度因子; K 的参数中包含 a_x 、 a_y 、 u_0 、 v_0 等相机内部参数, 所以被称为内参矩阵; M_1 中包含旋转矩阵 R 和平移向量 t , M_1 被称为外参矩阵, R 和 t 为外参; M 为投影矩阵。

至此, 计算机视觉系统的成像模型还是线性的, 即未发生畸变的理想状态。但是镜头加工工艺和相机位置会导致机器视觉成像系统成像过程中光线发生不同程度的弯曲, 造成图像畸变。因此引入图像非线性的畸变模型。非线性模型需要在相机坐标系转换到图像坐标系的过程中添加畸变, 即理想像素点坐标 (x_r, y_r) 添加畸变后转换为实际像素点坐标 (x_i, y_i) 。机器视觉系统的图像非线性畸变模型可表示为

$$\begin{cases} x_i = x_r + \sigma_x \\ y_i = y_r + \sigma_y \end{cases} \quad (7)$$

式中, σ_x 和 σ_y 是图像非线性畸变量, 包括径向、切向两种畸变。

径向畸变的数学模型可表示为

$$\begin{cases} \sigma_x = x_r(k_1 r^2 + k_2 r^4 + k_3 r^6) \\ \sigma_y = y_r(k_1 r^2 + k_2 r^4 + k_3 r^6) \end{cases} \quad (8)$$

式中, k_1 、 k_2 、 k_3 表示径向畸变参数, $r = \sqrt{x_r^2 + y_r^2}$ 。

切向畸变的数学模型可表示为

$$\begin{cases} \sigma_x = p_2(3x_r^2 + y_r^2) + 2p_1 x_r y_r \\ \sigma_y = p_1(x_r^2 + 3y_r^2) + 2p_2 x_r y_r \end{cases} \quad (9)$$

式中, p_1 、 p_2 为切向畸变参数。

将径向畸变和切向畸变的模型综合起来, 得到两种畸变综合作用于实际图像的非线性畸变模型:

$$\begin{cases} x_i = x_r + x_r(k_1 r^2 + k_2 r^4 + k_3 r^6) \\ \quad + p_2(3x_r^2 + y_r^2) + 2p_1 x_r y_r \\ y_i = y_r + y_r(k_1 r^2 + k_2 r^4 + k_3 r^6) \\ \quad + p_1(x_r^2 + 3y_r^2) + 2p_2 x_r y_r \end{cases} \quad (10)$$

式中, k_1 、 k_2 、 k_3 、 p_1 、 p_2 为畸变参数, 可表示成畸变矩阵 $D = (k_1, k_2, k_3, p_1, p_2)$; (x_r, y_r) 是理想的无畸变情况下的像素点坐标; (x_i, y_i) 是 (x_r, y_r) 畸变后的图像像素点

坐标^[8-10]。

2 数据集的获取

本文基于图像非线性模型模拟大量的机器视觉系统畸变程度不同的图像。由上文可知, 非线性模型中的畸变参数包括 k_1 、 k_2 、 k_3 、 p_1 、 p_2 , 所以数据集的生成主要是通过改变非线性模型中的畸变参数来随机生成畸变程度不同的图像。因此, 基于卷积神经网络的回归功能, 数据集里的畸变图像就无需分类处理。实验中, 本文利用 MATLAB 以及 LABEL 函数记录畸变图像所对应的畸变参数矩阵 D 。在机器视觉系统的实际工作中, 采用定焦拍摄, 且常用的焦距是 4~300 mm 范围内的许多等级。将标准胸环靶面图像作为被摄对象, 改变 k_1 、 k_2 、 k_3 、 p_1 、 p_2 的赋值, 在常用焦距范围内改变焦距 f , 使产生畸变的图像在靶面上成像, 进而得到可用于训练的畸变图像。用来生成畸变图像数据集的胸环靶面灰度图像如图 3 所示。

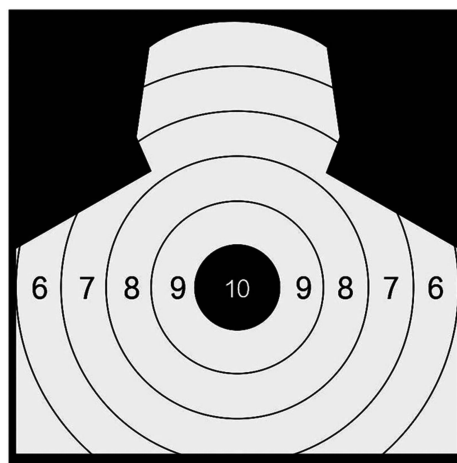


图 3 未产生畸变的胸环靶面图像

在对畸变图像模拟的过程中, 不断地对畸变参数随机赋值, 并在机器视觉系统的焦距范围内不断调节焦距, 使图像成像在靶面上。对 5 个畸变参数赋值时, 要进行协调赋值。如果都在最大值范围内进行赋值, 那么图像变形会非常严重, 将无法使用。最后经过不断模拟发现, 在一定范围内分别对畸变参数进行随机赋值, 生成图像的可用度更高。具体范围如下:

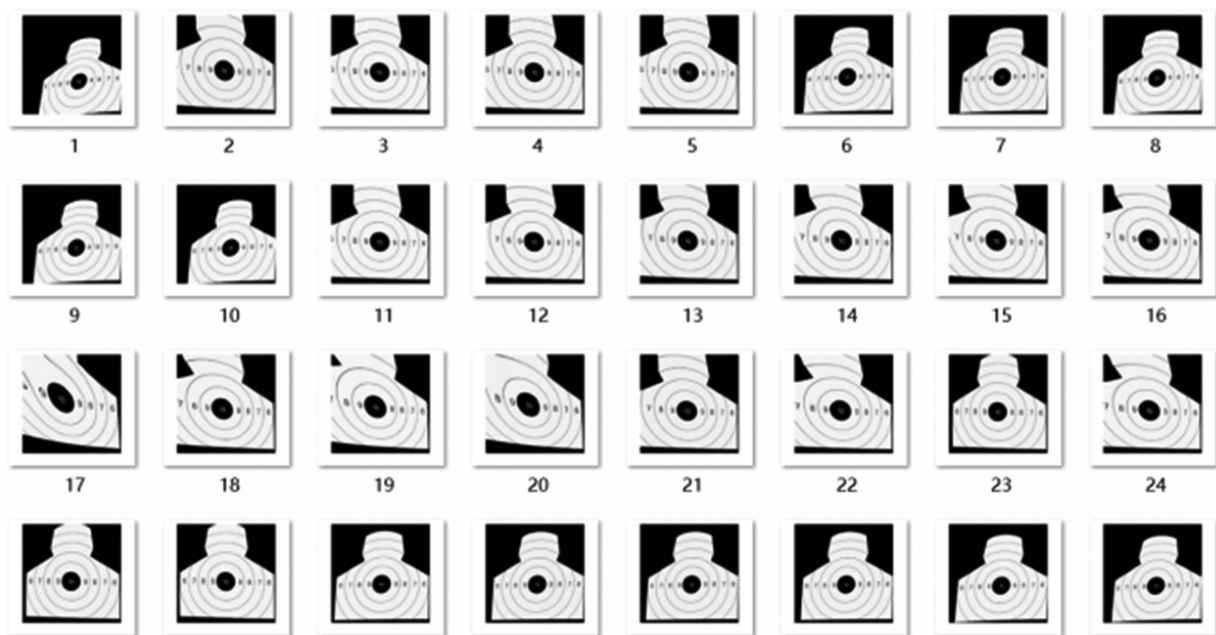


图 4 部分畸变图像示例

$$k_1, k_2, k_3, p_1, p_2 \in E-01 * [-1, 1] \cup E-02 * [-1, 1] \cup E-03 * [-1, 1] \dots \dots \cup E-09 * [-1, 1] \quad (11)$$

本文基于图像非线性畸变模型模拟生成了 12000 张畸变图像。数据集里训练集的部分畸变图像示例如图 4 所示。

3 网络模型设计及实验

3.1 模型设计

卷积神经网络一般由输入层、隐藏层和输出层组成。其中，隐藏层通常由卷积层、池化层和全连接层构成。卷积神经网络模型的性能是否优良主要取决于网络模型中各层的数量以及各层参数的选取配置。模型层数越多，其非线性拟合能力越强，模型的运算复杂度越高^[11-14]。本文利用卷积神经网络对畸变图像进行回归，并对图像非线性畸变模型进行求解，得到其中的畸变参数。

本文设计的用于图像畸变校正算法的卷积神经网络结构由输入层、三层卷积层、三层池化层、两层全连接层以及输出层组成。本文的训练模型的卷积层采用小卷积核。通过叠加多层卷积层，使特征图像可经过更多的非线性映射处理，提高了网络训练中算法的拟合能力，

并提升了模型性能。此外，采用小卷积核可以减少网络参数，计算速度更快。对于网络结构中的每个卷积层，本文都采用了 relu 激活函数，并在其后应用了最大池化层。在这一系列层级之后，对最后一层池化层的输出结果进行扁平化处理并将其传入全连接层。最后的输出结果分别为每个畸变参数。卷积神经网络的整体模型结构如图 5 所示。

3.2 畸变校正

选择三张畸变的胸环靶面图像(见图 6)，将其输入到已经训练好的模型中。得到的输出结果为 k_1 、 k_2 、 p_1 、 p_2 、 k_3 等 5 个图像非线性畸变模型畸变参数值，即每个系统的畸变参数。对这三个机器视觉系统的图像非线性畸变模型进行求解，得到模型预测的畸变参数： $k_1 = -1.223522 \times 10^{-2}$ ， $k_2 = 3.572984 \times 10^{-3}$ ， $p_1 = -6.985934 \times 10^{-6}$ ， $p_2 = 9.065283 \times 10^{-4}$ ， $k_3 = -2.906040 \times 10^{-7}$ ； $k_1 = 3.982671 \times 10^{-4}$ ， $k_2 = -5.925159 \times 10^{-4}$ ， $p_1 = -1.168219 \times 10^{-4}$ ， $p_2 = -1.075688 \times 10^{-5}$ ， $k_3 = -6.867518 \times 10^{-5}$ ； $k_1 = -7.136880 \times 10^{-4}$ ， $k_2 = 1.187411 \times 10^{-4}$ ， $p_1 = -9.908408 \times 10^{-5}$ ， $p_2 = 5.333640 \times 10^{-2}$ ， $k_3 = 6.974185 \times 10^{-3}$ 。以上参数分别为这三个机器视觉系统的畸变参数。

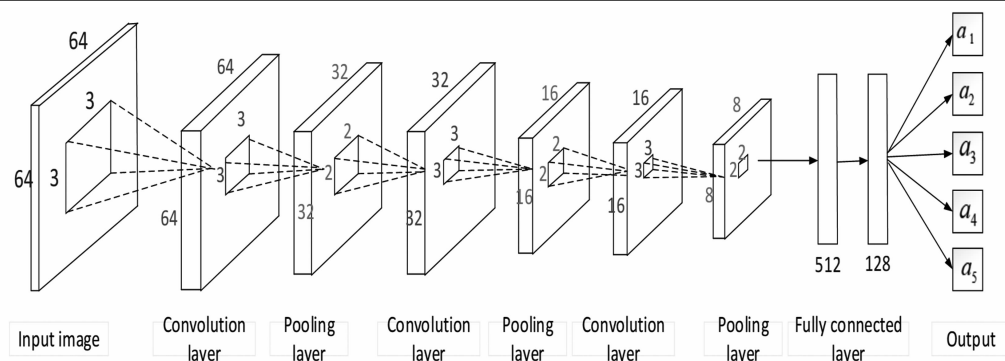


图 5 卷积神经网络的整体模型结构图

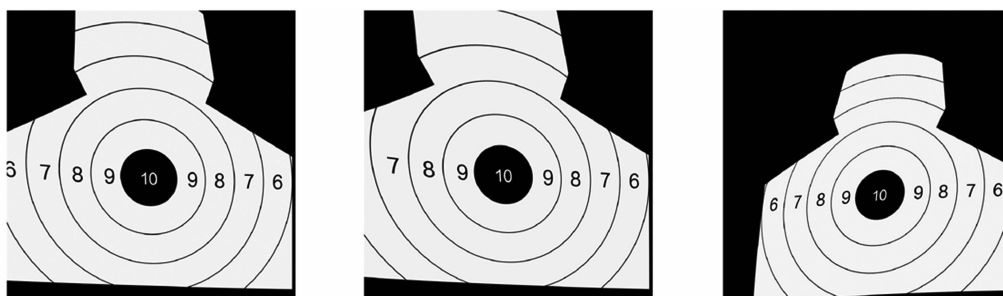


图 6 棋盘格畸变图像

然后根据上述畸变参数, 分别利用系统的畸变模型对其所测工件的畸变图像进行校正。本文在 MATLAB 中设计了用于图像校正的程序。图 7 为该程序的流程图。

校正开始时, 分别输入畸变的工件图像。先将图像中各像素的灰度值依次按坐标提取为数组, 然后基于已求得的图像畸变模型依次求取每个像素点的实际坐标点 (x_r, y_r) 所对应的畸变图像坐标点 (x_i, y_i) 。如果计算出的 (x_i, y_i) 点不是整数, 则用二次线性插值法计算该点的灰度值, 并用坐标点 (x_i, y_i) 的灰度来填充校正后图像上的 (x_r, y_r) 点。最后将存储校正后各像素灰度值的数组恢复为图像, 得到校正后的工件图像。图像校正前后的对比效果如图 8 所示。

采用传统四点透射方法的校正效果如图 9 所示。通过对比图 8 与图 9 可以看出, 本文针对胸环靶面图像提出的畸变校正算法效果较好, 有利于提升自动报靶系统的报靶精度。

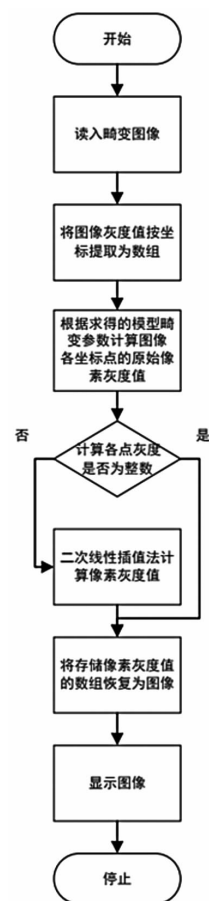


图 7 程序流程图

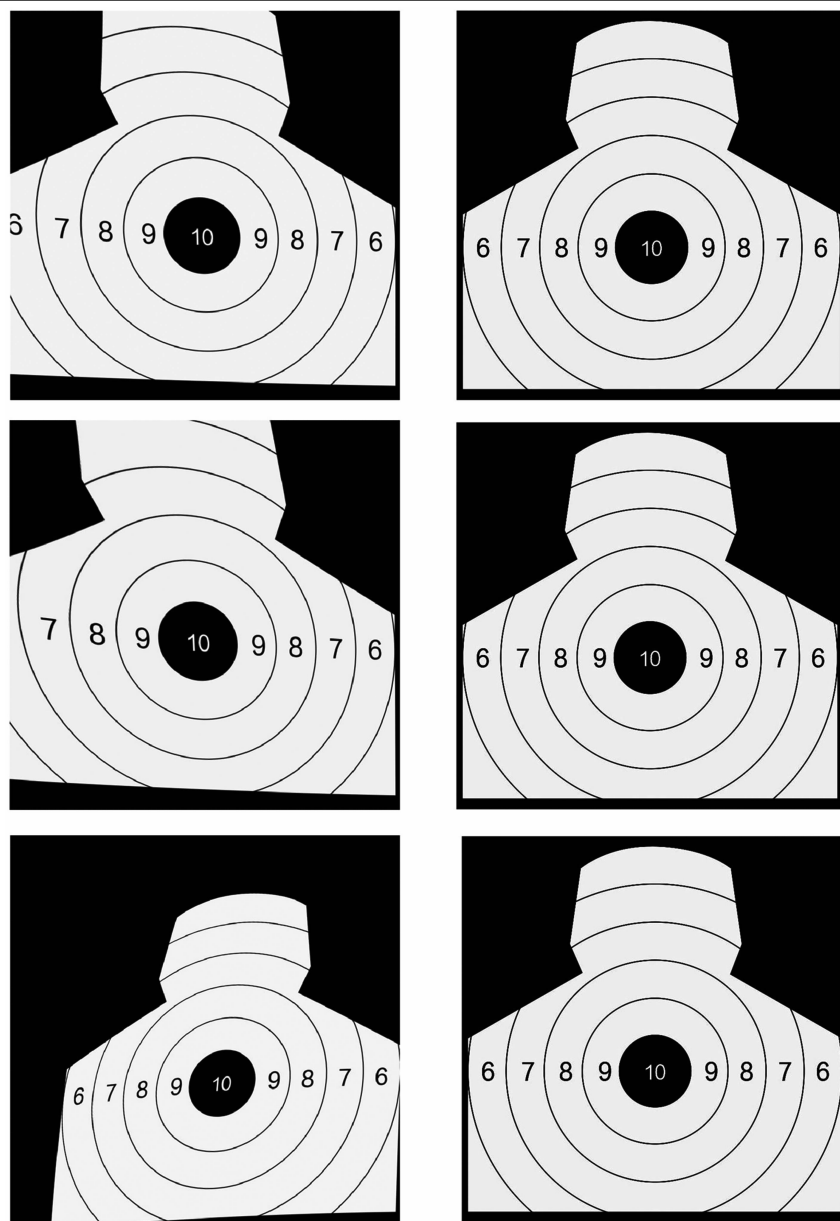


图 8 本文方法得到的图像

4 结束语

本文通过对传统自动报靶系统的研究分析,针对图像畸变影响报靶精度的问题,提出了一种基于卷积神经网络的畸变校正算法。采用对图像畸变模型的畸变参数进行后求解的方法,完成了对系统图像的校正。传统的算法在求解畸变参数之前,需要先计算计算机视觉系统结构的其他参数(如相机焦距、像素尺寸因子等),而且需要多张多角度的畸变图像。畸变参数的求解过程较为繁琐,给图像畸变校正带来不便。而本文算法只需要一张畸变模板图

像就能得出图像非线性模型的畸变参数。我们利用 MATLAB 模拟出大量机器视觉系统畸变程度不同的图像,并将其作为数据集对神经网络进行训练。训练完成后,输入一张畸变图像就可以得到该图像的畸变参数,并利用该参数完成对图像的畸变校正。校正后的图像更能满足计算机视觉系统的使用需求。本文利用不同的胸环靶面畸变图像对该算法进行了验证。与传统校正算法的对比结果表明,本文提出的校正算法效果更优,有利于提高基于计算机视觉的自动报靶系统的报靶精度。

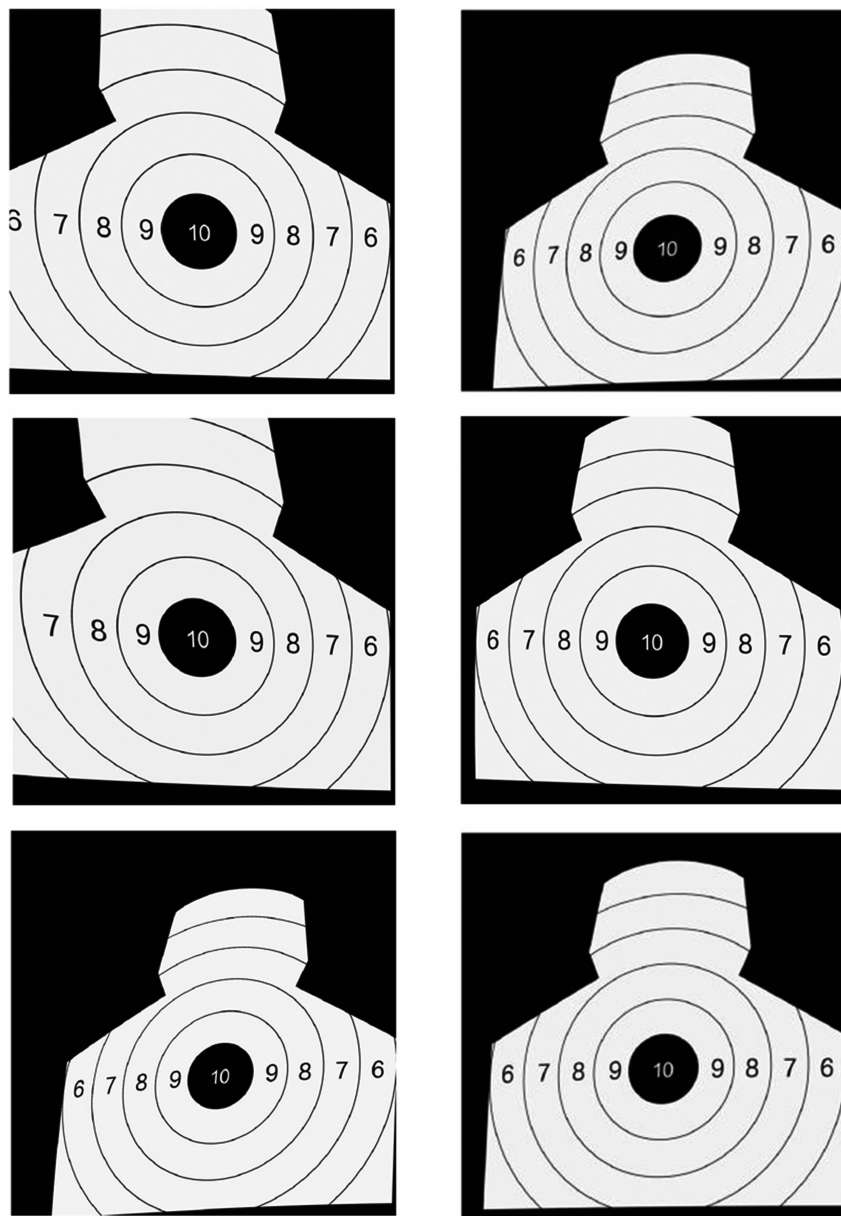


图9 传统四点透射方法的校正效果

参考文献

- [1] 黄应清, 陈晓明, 谢志宏, 等. 一种基于计算机视觉的胸环靶弹孔检测方法 [J]. 兵器装备工程学报, 2021, 42(11): 157-162.
- [2] Shao Y F, Zhang L, Wu G W. The visual detection of bullet holes in a target [C]. Beijing: 5th International Conference on Signal Processing, 2000.
- [3] Sun Q, Li C, Zheng J. Effects of camera lens distortion on particle size and shape characterizations [J]. *Geotechnical Testing Journal*, 2021, 44(4): 929-947.
- [4] 杨丽凤, 武利生, 刘建红, 等. 机器视觉尺寸检测系统中安装误差分析 [J]. 太原理工大学学报, 2004, 48(3): 332-334.
- [5] 李梦祺. 基于视觉检测的胸环靶自动报靶系统研究 [D]. 沈阳: 沈阳工业大学, 2018.
- [6] 梁伟, 杨家巍, 崔健. 基于视频的实弹和激光双模报靶系统及报靶方法: CN113983871B [P]. 2023-04-25.
- [7] 刘文韬, 张燕, 徐雅楠, 等. 基于图像处理的自动报靶方法及系统: CN115690408A [P]. 2023-02-03.

- [8] 张志红, 余国河, 田东哲, 等. 相机标定对玉米自动化观测评估结果的影响 [J]. *单片机与嵌入式系统应用*, 2019, **19**(4): 26–29.
- [9] 吴福朝, 阮宗才, 胡占义. 非线性模型下的摄像机自标定 [J]. *计算机学报*, 2002, **25**(3): 276–283.
- [10] 杨雪. 基于 CNN 的机器视觉系统图像畸变校正算法研究 [D]. 长春: 吉林大学, 2022.
- [11] Li Z, Zhao X. BP artificial neural network based wave front correction for sensor-less free space optics communication [J]. *Optics Communications*, 2017, **385**: 219–228.
- [12] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, **521**(7553): 436–444.
- [13] Samudre P, Shende P, Jaiswal V. Optimizing performance of convolutional neural network using computing technique [C]. Pune: 2019 IEEE 5th International Conference for Convergence in Technology, 2019.
- [14] Cresson R. A Framework for Remote Sensing Images Processing Using Deep Learning Techniques [J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, **16**(1): 25–29.
- [15] Pei Y, Huang Y, Zou Q, et al. Effects of image degradation and degradation removal to CNN-based image classification [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, **43**(4): 1239–1253.