专题 农业信息高质量图谱数据

奶牛个体头部多模态图像深度学习训练数据集

杜林煊[1,2]，曹姗姗[1,2]，刘婷婷[1,2]，孙伟[1,2*]，孔繁涛[3*]

1. 中国农业科学院农业信息研究所，北京 100081
2. 国家农业科学数据中心，北京 100081
3. 中国农业科学院农业经济与发展研究所，北京 100081

摘要：随着农业现代化与精准农业的推进，畜牧业的智能化与精准化管理在动物健康监测、养殖过程管理以及精准养殖等领域展现出广泛的应用前景。基于图像分析的非接触式健康监测技术，利用计算机视觉与红外成像，可实时监控奶牛的生理状态和行为特征，为疾病预警与生产管理优化提供支持。本研究构建了一个涵盖奶牛个体头部多模态图像的数据集，包含 50 头奶牛在自然活动状态下，不同光照条件（昼间和夜间）下采集的图像数据。本数据集包括 2636 组经过配准的可见光图像（VIS）和红外热成像图像（IRT），并对 IRT 图像中的奶牛头部区域进行了精确标注。通过模糊检测、图像质量评估和数据增强等技术优化了数据质量，确保数据多样性和真实性。本数据集可为跨模态学习、图像配对与行为分析研究提供基础，助力智能监测系统与精准农业的发展。

关键词：奶牛个体检测；红外测温图像；多模态数据；精准农业；奶牛健康监测；深度学习

数据库（集）基本信息简介

数据库（集）名称	奶牛个体头部多模态图像深度学习训练数据集
数据作者	孔繁涛，孙伟
数据通信作者	孙伟（sunwei02@caas.cn）
数据时间范围	2024年
地理区域	北京
数据量	1.31 GB
数据格式	*.png,*.txt
数据服务系统网址	https://doi.org/10.57760/sciencedb.agriculture.00184
基金项目	国家重点研发计划（2024YFD2000305）；中国农业科学院科技创新工程（10-IAED-RC-09-2024）；新疆生产建设兵团农业关键核心技术攻关项目（NYHXGG2023AA310-06）。
数据库（集）组成	数据集共包括9个数据文件，其中：train、test、val分别为训练集测试集和验证集，其中每个文件分别包含img、VIS、lable三个文件夹，分别存储红外测温图像，可见光图像和图像标签。

www.sciengine.com/CSD
ISSN 2096-2223
CN 11-6035/N

文献 CSTR：
32001.14.11-6035.nasdc.2025.0001.zh

文献 DOI：
10.11922/11-6035.nasdc.2025.0001.zh
数据 DOI：
10.57760/sciencedb.agriculture.00184
文献分类：农学

收稿日期：2024-12-24
开放同评：2024-12-25
录用日期：2024-12-30
发表日期：2025-02-12

* 论文通信作者
孔繁涛: kongfantao@caas.cn
孙 伟: sunwei02@caas.cn

引　言

随着中国农业现代化进程的不断推进，以及乡村振兴战略深入实施，畜牧业发展取得一定成效，畜牧业产业化、集约化和组织化程度不断提高，现代化特征逐步显现[1]。根据国家统计局（2023）的数据，2023 年中国牛存栏量达 10508.51 万头，同比较 2022 年增长了 292.66 万头，增幅为 2.78%。与 2013 年的 8985.76 万头相比，近 10 年间增量达 1522.75 万头、增幅为 14.49%[2]。畜牧业的智能化与精准化管理逐渐成为提升生产效率、开展动物健康监测以及实现精准养殖的重要方向。尽管如此，奶牛监测方法依赖人工、效率低且易受环境影响，难以实现高效、精准、实时的监控，亟需新型技术手段来提高监测能力[3]。近年来，随着计算机视觉和红外成像技术的迅猛发展，关于图像分析的动物健康监测技术已在疾病预警、行为监控等方面得到了广泛应用[4-5]。为了提高养殖业生产效率并减少人为误差，基于图像处理和传感技术的非接触式监测手段逐渐得到了广泛应用[6]。其中，奶牛的可见光图像和红外测温图像作为重要检测数据，可综合提供外部状态与生理温度信息，有助于疾病早期发现、生长监测及饲养管理优化。

本研究构建一个可见光图片与红外测温图像（IRT）对应的奶牛个体头部多模态图像数据集，为跨模态学习、多模态图像配准与融合和动物健康监测等研究提供了数据支持。可见光图像与红外图像的结合，为多种实际任务提供了更丰富的视觉信息，能够有效提升系统的感知能力与应用效能。

目前，医学成像和安全监控等领域已有包括红外图像和可见光图像的多模态数据集[7-8]，已有部分数据集涉及农业生产和牲畜健康监测领域[9]，但是在奶牛的可见光图像与红外测温图像方面，相关数据集较少。本数据集弥补了这一不足，提供了可见光图像和红外图像对应的奶牛图像数据。

本数据集包含一系列奶牛的可见光图像与红外测温图像，每一组图像都经过对齐，确保红外图像中的温度分布与可见光图像中的视觉信息具有高度一致性。基于本数据集，研究人员可以在可见光图像与红外测温图像之间开展更为深入的跨模态学习与图像配对任务。

红外图像能够准确反映动物体表温度变化，可见光图像则提供丰富的外部状态信息。两者结合可用于配对与融合任务，同时为跨模态学习、图像检索和目标检测等研究提供数据支持。

1　数据采集和处理方法

1.1　数据采集方法

实验数据采集于中国北京市昌平区某养牛厂，实验对象为两个分区共计 50 头奶牛，拍摄时间为 2024 年 9 月。本次实验所用红外摄像机为高德红外 IPN384，进行高精度测温，测温精度±0.5℃，热灵敏度≤50mK@30℃。可见光相机为 HISKKVISION 变焦筒形网络摄像机 DS-2XA2668F-LZS 图像分辨率为 1080p。

摄像机固定于移动平台，高度距地面 1.5m，摄像机距离牛 1.5-2 m 处于红外测温最佳距离范围内。实时调整摄像机角度，确保可以拍摄到奶牛的头部，同时启动可见光相机和红外相机进行视频录制。视频录制过程中，所有奶牛都在自然活动状态下进行采集，确保数据的真实性和应用场景的相符性，见图 1 图像采集示意图。



图 1　图像采集示意图

Figure 1　Diagram of image acquisition

1.2　数据处理

对采集到的多段奶牛视频进行时间轴对齐和剪辑，筛除了不含牛只目标的画面片段，得到预处理视频数据，帧速率 25 帧每秒。在本研究中，我们采用多种图像处理技术对可见光（VIS，Visible Spectrum）和红外（IRT，Infrared Thermography）视频帧进行预处理，以提高数据质量并确保后续分析的准确性。处理过程包括图像正畸、模糊检测、相似筛选、图像质量评估、剪裁等操作，具体步骤如图 2 所示。

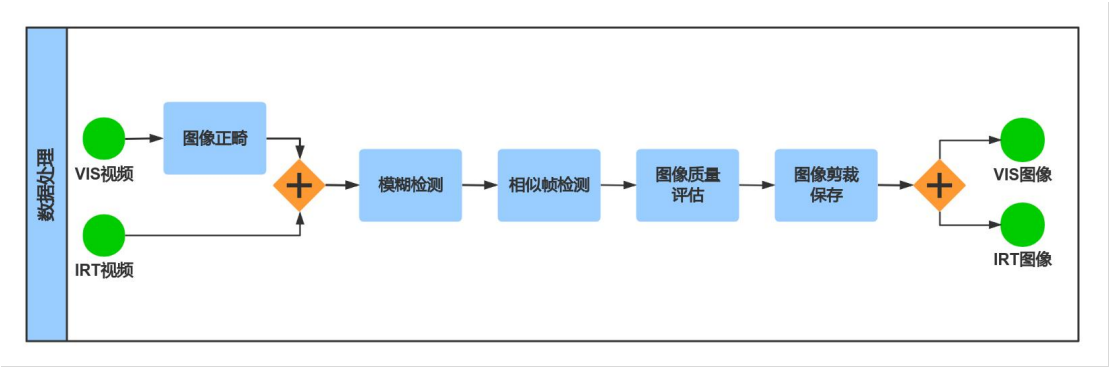


图 2　数据处理流程图

Figure 2　Data processing flowchart

为了消除因镜头特性或拍摄条件导致的图像几何畸变，使图像更接近真实场景，进行图像正畸，首先通过基于棋盘格图案的相机标定来获取相机的内参矩阵和畸变系数，计算提取角点信息，得到相机的内参矩阵和畸变系数，并利用该参数对每一帧图像进行矫正。计算最佳相机矩阵来确定裁剪区域，以去除图像边缘的无效区域，确保分析区域集中在目标物体上。

对一组图片进行对准后，根据 IRT 图像计算感兴趣区域（ROI，Region of Interest）对 VIS 图片进行剪裁，并将剪裁后的 VIS 图像缩放至 IRT 图像相同（512px × 384px），以确保所有图像具有统一的尺寸和分辨率，便于后续分析。图为处理好的 VIS 和 IRT 图像。图像处理结果如图 3 所示，其中图 3（a）为 VIS 的原图像，图 3（b）为经过图像正畸后的图像结果，图 3（c）为根据 ROI 剪裁后的 VIS 图像，图 3（d）为对应的 IRT 图像。



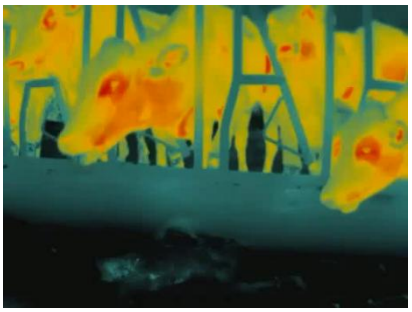
（a）原 VIS 图像



（b）图像正畸结果



（c）剪裁后的 VIS 图像



（d）对应的 IRT 图像

图 3　VIS 和 IRT 数据处理结果对比图

Figure 3　Comparison of VIS and IRT data processing results

对数据集进行数据增强，增加训练数据的多样性，对图像进行旋转、平移、缩放、裁剪、翻转等变换，如表 1 数据增强参数表所示。并对标注文件中的坐标点进行相应的转换，同时只对训练集进行数据增强操作，每张图像随机增强 3 次，如图 4 所示。其中图 4（a）为 IRT 原图，图 4（b）、（c）、（d）为数据增强后的结果。图像文件被随机打乱，然后根据设定的比例 80%用于训练，10%用于验证，10%用于测试，划分为训练集（Train Se）t、验证集（Validation Set）和测试集（Test Set）三个子集。



（a）原 IRT 图



（b）随机旋转和放大



（c）随机平移、旋转和剪裁



（d）随机旋转、缩小、剪裁和翻转

图 4　一组数据增强结果

Figure 4　One set of data Augmentation results

表 1　数据增强参数表

Table 1　Table of data augmentation parameters

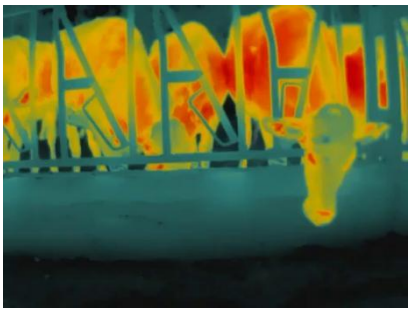
序号	操作	描述	参数值
1	旋转	随机旋转图像一定的角度。	旋转角度：(-15°, 15°) 旋转中心：(width // 2, height // 2)
2	平移	随机平移图像。	水平平移范围：(-50px, 50px) 垂直平移范围：(-50px, 50px)
3	缩放	随机缩放图像大小。	缩放因子范围：(0.8, 1.2)
4	水平翻转	随机对图像进行水平翻转。	是否翻转：True 或 False（随机选择）
5	裁剪	随机裁剪图像的一部分。	裁剪比例：0.9

2　数据样本描述

数据为对照的 VIS 和 IRT 图像组，分别存放在两个文件夹中，对应图像的文件名相同，如图 5 数据示例对比图所示，典型图像分为两个部分，图 5（a）（b）分别为夜间采集的 VIS 图像以及对应的 IRT 图像，图 5（c）（d）分别为昼间采集的 VIS 图像以及对应的 IRT 图像。



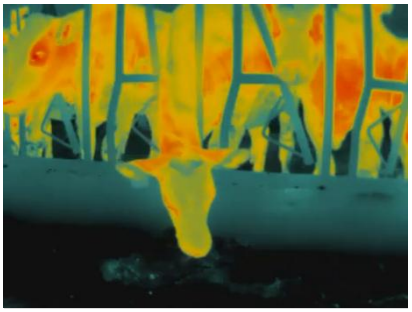
（a）夜间 VIS 图像



（b）夜间 IRT 图像



（c）昼间 VIS 图像



（d）昼间 IRT 图像

图 5　数据示例对比图

Figure 5　Comparison of data samples

原始数据集共包含 759 组图像，其中夜间图像 397 组，昼间图像 362 组。经过数据增强后的数据共 2636 组。针对红外测温图像中的奶牛头部区域，采用 Supervisely 平台进行标注，标注方式为使用矩形框标记奶牛头部，标签命名为“head”。以下图像标注示例如图 6 所示。

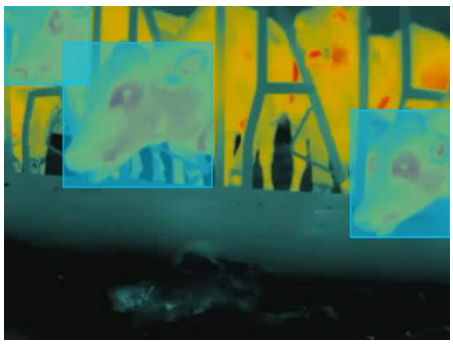


图 6　数据标注示例图

Figure 6　Diagram of data annotation samples

标注结果以 YOLO 标签格式存储，每张图像对应一个同名的 TXT 文件。其中类别编号 0 为标签“head”的类别编号，中心点 x 坐标，中心点 y 坐标，宽度，高度，分别表示标注框中心点在图像中的坐标以及标注框的大小，并将参数值归一化，取值范围为[0, 1]。数据标签示例如表 2 所示。

表 2　数据标签示例表

Table 2　Table of data annotation examples

序号	类别编号	中心点 x 坐标	中心点 y 坐标	宽度	高度
1	0	0.0957	0.1159	0.1875	0.2318
2	0	0.2988	0.3242	0.332	0.4297
3	0	0.8877	0.5	0.2207	0.375

3　数据质量控制和评估

标准化采集条件置于统一的光照条件、温度、拍摄角度等，确保在固定的环境条件下进行图像采集，减少外部干扰，避免因环境变化而导致数据偏差。IRT 和 VIS 图像在同一场景下拍摄，且固定位置、角度对齐，以便后续对比分析。利用几何校正方法对图像进行对齐。

在图像整理筛查过程中，通过模糊检测、图像质量评估、相似筛选来进行数据的质量控制。本数据集使用基于拉普拉斯算子（Laplacian）的图像清晰度评估方法[10]。通过计算图像的拉普拉斯方差，若方差低于设定的阈值，则认为该帧图像为模糊图像，予以剔除。使用帧差法和结构相似性指数（SSIM）评估两组图像之间的视觉相似性，当 SSIM 值高于预设阈值（0.95）且帧差值低于阈值（1.0）时，认为帧之间的差异不显著，剔除差异不显著的图像以减少重复数据和无效帧。对于红外图像，通过 Canny 边缘检测进一步进行质量评估，提取图像轮廓，并计算轮廓数量和噪声水平，以剔除模糊和图像噪声过大的图像。最终进行手动检查，以确保数据质量，对总帧数 11923 的图像进行筛选，具体的各项图像质量控制方式及对应的数据处理数量，详见表 3。

表 3　数据质量控制表

Table 3　Table of data quality control

序号	项目		数量	占比
1	质量控制方式	相似帧	1004	8.42%
2		可见光模糊	1469	12.32%

序号	项目		数量	占比
3		红外噪声高	8623	72.32%
4		手动检查	68	0.57%
5	保存帧数		759	6.37%
6	总删除帧数		11164	93.63%

对每一帧图像，在确保其质量合格后，进行图像标注工作。标签空间热图见图 7 所示，显示所有类别对象的平均空间位置和位置，突出显示最常见和罕见的对象位置，能够观察到标签基本均匀分布在整个图像区域，边缘分布较少，符合图像正常标签位置，没有明显的异常聚集。

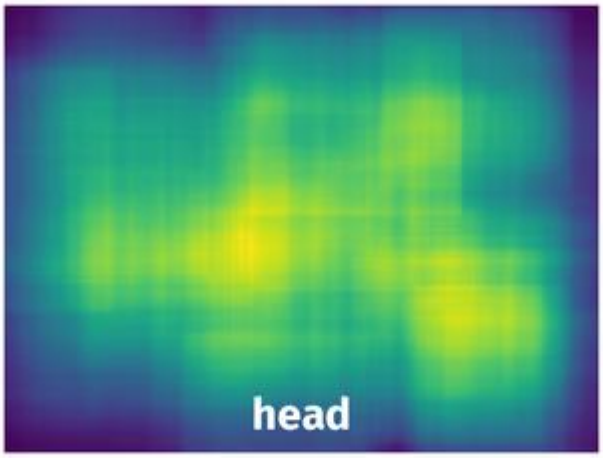


图 7　标签空间热图

Figure 7　Heatmap of the label space

4　数据价值

与国内外现有的数据集相比，本数据集重点关注奶牛头部的可见光图像与红外测温图像对照数据，能够进行多模态数据融合研究。数据集的潜在应用领域广泛，在动物健康监测和农业自动化领域具有应用价值。红外图像能够实时反映奶牛体表温度变化，而可见光图像则提供了丰富的视觉信息，二者结合能够更准确地评估动物的健康状态。

通过对多模态数据的分析与融合，可以更好地理解奶牛的生理特征、疾病症状以及行为规律，用来训练基于机器学习或深度学习的智能识别系统。为智能监测系统、疾病预警系统等农业科技创新提供数据支持，推动精准农业和智能养殖的发展，具有重要的研究和应用意义。

数据作者分工职责

杜林煊（2000—），男，山西运城人，硕士，研究方向为农业信息化。主要承担工作：数据采集和数据论文撰写。

曹姗姗（1984—），女，黑龙江哈尔滨人，博士，副研究员，研究方向为农林时空信息智能分析。主要承担工作：数据汇总整理。

刘婷婷（1985—），女，北京人，,硕士，助理研究员，研究方向为科学数据管理。主要承担工作：数据审核。

孙伟（1978—），男，山东海阳人，博士，研究员，研究方向为农林时空信息智能分析。主要承担工作：数据质量控制与综合管理。

孔繁涛（1968—），男，山东滕州人，博士，研究员，研究方向为农业信息技术。主要承担工作：研究思路设计与数据审核。

参考文献

[1] 于杰, 王明利. 中国畜牧业现代化研究进展与未来展望[J]. 中国农业资源与区划, 1–10. [YU J, WANG M L. Research Progress and Future Prospects of Modernization in China's Animal Husbandry [J]. China Agricultural Resources and Regional Planning, 1–10.]

[2] 国家统计局. 中国统计年鉴 2023[M]. 北京: 中国统计出版社, 2023. [NATIONAL BUREAU OF STATISTICS. China Statistical Yearbook 2023 [M]. Beijing: China Statistics Press, 2023.]

[3] 焦杰, 齐咏生, 刘利强, 等. 一种场景自适应的双分支牛脸高效识别算法[J]. 电子学报, 2024, 52(09): 3251–3261. [JIAO J, QI Y S, LIU L Q, et al. A Scene-Adaptive Dual-Branch Efficient Cow Face Recognition Algorithm [J]. Acta Electronica Sinica, 2024, 52(09): 3251–3261.]

[4] WANG Y, KANG X, CHU M, et al. Deep learning-based automatic dairy cow ocular surface temperature detection from thermal images [J]. Computers and Electronics in Agriculture, 2022, 202: 107429.

[5] JADDOA M A, GONZALEZ L, CUTHBERTSON H, et al. Multiview eye localisation to measure cattle body temperature based on automated thermal image processing and computer vision [J]. Infrared Physics & Technology, 2021, 119: 103932.

[6] ZHANG B, XIAO D, LIU J, et al. Pig eye area temperature extraction algorithm based on registered images [J]. Computers and Electronics in Agriculture, 2024, 217: 108549.

[7] 黄家才, 常国卫, 洪颖, 等. 基于事件帧和RGB帧融合的交通场景目标检测方法[J]. 计算机工程与应用, 1–13. [HUANG J C, CHANG G W, HONG Y, et al. Target Detection Method for Traffic Scenes Based on Event Frames and RGB Frame Fusion [J]. Computer Engineering and Applications, 1–13.]

[8] 辛家隆, 张雪松, 强彦, 等. 基于改进焦点损失与图像特征融合的CT图像分类方法[J]. 太原理工大学学报, 1–14. [XIN J L, ZHANG X S, QIANG Y, et al. CT Image Classification Method Based on Improved Focal Loss and Image Feature Fusion [J]. Journal of Taiyuan University of Technology, 1–14.]

[9] 黄小平. 奶牛体况评分目标检测数据集[M]. 第二版. 2024. [HUANG X P. Cow Body Condition Scoring Object Detection Dataset [M]. V2 ed., 2024.]

[10] ZHU M, YU L, WANG Z, et al. A survey on objective evaluation of image sharpness [J]. Applied Sciences, 2023, 13(4): 2652.

论文引用格式

杜林煊，曹姗姗，刘婷婷，等. 奶牛个体头部多模态图像深度学习训练数据集[J/OL]. 中国科学数据，2025, 10(1). (2025-02-12). DOI: 10.11922/11-6035.nasdc.2025.0001.zh.

数据引用格式

孔繁涛，孙伟. 奶牛个体头部多模态图像深度学习训练数据集[DS/OL]. V2. Science Data Bank, 2024. (2025-02-12). DOI: 10.57760/sciencedb.agriculture.00184.

A dataset of deep learning training for multimodal head images of individual cows

DU Linxuan[1,2], CAO Shanshan[1,2], LIU Tingting[1,2], SUN Wei[1,2*], KONG Fantao[3*]

1. Agricultural Information Institute of CAAS, Beijing 100081, P. R. China
2. National Agriculture Science Data Center, Beijing 100081, P. R. China
3. Institute of Agricultural Economics and Development of CAAS, Beijing 100081, P. R. China

* Email: kongfantao@caas.cn (KONG Fantao) ; sunwei02@caas.cn (SUN Wei)

Abstract: With the advancement of agricultural modernization and precision agriculture, the intelligent and precision management of livestock farming has shown great potential for applications in areas such as animal health monitoring, farming process management, and precision farming. Leveraging computer vision and infrared imaging, image analysis-based non-contact health monitoring technologies enable real-time monitoring of cows' physiological states and behavioral characteristics, which can support early disease detection and production management optimization. This study develops a dataset of multimodal head images of individual cows, comprising images collected from 50 cows under natural activity conditions and varying lighting environments (daytime and nighttime). The dataset includes 2,636 pairs of co-registered visible light (VIS) images and infrared thermographic (IRT) images, with precise annotations of cow head regions in the IRT images. Techniques such as blur detection, image quality assessment, and data augmentation were employed to optimize data quality, ensuring diversity and authenticity. This dataset can serve as a foundation for research on cross-modal learning, image pairing, and behavioral analysis, contributing to the development of intelligent monitoring systems and precision agriculture.

Keywords: individual cow detection; infrared thermography; multimodal data; precision agriculture; cow health monitoring; deep learning

Dataset Profile

Title	A dataset of deep learning training for multimodal head images of individual cows
Data corresponding author	SUN Wei (sunwei02@caas.cn)

Data authors	KONG Fantao, SUN Wei
Time range	2024
Geographical scope	Beijing
Data volume	1.31 GB
Data format	*.png,*.txt
Data service system	<https://doi.org/10.57760/sciencedb.agriculture.00184>
Sources of funding	National Key Research and Development Program of China (2024YFD2000305); Agricultural Science and Technology Innovation Project of Chinese Academy of Agricultural Sciences (10-IAED-RC-09-2024); Key Core Agricultural Technology Research Project of the Xinjiang Production and Construction Corps (NYHXGG2023AA310-06).
Dataset composition	The dataset consists of a total of 9 data files, including three sets (train, test, and val for training, testing, and validation, respectively). Each of these files contains three folders: img, VIS, and label, which store the infrared thermography images, visible light images, and image labels, respectively.