

研究论文 ORIGINAL PAPERS

基于 DenseFuse 网络的无人机载 红外和可见光鹿科动物图像融合

李汶佼¹, 包衡², 杜化林³, 李洋⁴,
张卫华⁴, 杨琨⁴, 马光凯^{1,2*}, 姜广顺^{2*}

(1. 东北林业大学计算机与控制工程学院, 哈尔滨 150040;

2. 国家林业和草原局猫科动物研究中心, 东北林业大学野生动物与自然保护地学院, 哈尔滨 150040;

3. 内蒙古森工集团, 牙克石 022150;

4. 内蒙古大兴安岭汗马国家级自然保护区管理局, 根河 022359)

稿件运行过程

收稿日期: 2024-05-14
修回日期: 2024-05-17



摘要

关键词: 红外图像;
可见光图像;
野生动物;
图像融合

Keywords: Infrared image;
Visible image;
Wildlife;
Image fusion

中图分类号: Q958.1; TP391.4

文献标志码: A

文章编号:

2310-1490(2025)-03-0514-09

DOI: 10.12375/ysdwxb.20250305

野生鹿科(Cervidae)动物作为生态系统的组成部分,在维持生态平衡方面扮演着至关重要的角色。无人机影像技术在野生动物监测中的应用已日趋成熟,但受自然光照条件和野外环境复杂多变的影响,单一光谱成像技术很难得到高质量的野生鹿科动物图像。因此提出一种基于 DenseFuse 网络的图像融合算法,通过无人机搭载的多光谱成像设备,将红外图像与可见光图像融合,同时保留红外图像的轮廓信息和可见光图像的外貌信息,提高了监测图像质量。基于野生鹿科动物影像数据集,采用多种图像融合策略进行实验,对红外图像与可见光图像的融合效果展开细致对比。结果显示,通过使用 l_1 -norm 融合策略所获得的综合评价指标最优,经该策略融合后得到的图像平均信息熵达到了 6.965。这一结果表明,本研究所提出的无人机多光源图像融合算法能够为野生动物监测工作提供可靠的技术支撑。

基金项目: 国家重点研发计划项目(2023YFF1305000); 中央高校基本科研业务费专项基金项目(2572022DS04)

第一作者简介: 李汶佼(1998—),女,硕士研究生;主要从事目标检测、图像分类研究。E-mail: wenjiaoli_hrb@163.com

* 通信作者: 马光凯, E-mail: gkmahr@163.com; 姜广顺, E-mail: jgshun@126.com

Unmanned Aerial Vehicle Equipped with both Infrared and Visible for Cervidae Image Fusion Based on DenseFuse Network

LI Wenjiao¹, BAO Heng², DU Hualin³, LI Yang⁴,
ZHANG Weihua⁴, YANG Kun⁴, MA Guangkai^{1,2*}, JIANG Guangshun^{2*}

(1. College of Computer and Control Engineering, Northeast Forestry University, Harbin 150040, China;

2. Feline Research Center of National Forestry and Grassland Administration,

College of Wildlife and Protected Area, Northeast Forestry University, Harbin 150040, China;

3. Inner Mongolia Forestry Industry Group, Yakeshi 022150, China;

4. Inner Mongolia Hanma National Nature Reserve, Genhe 022359, China)

Abstract: As an integral part of the ecosystem, wild Cervidae animals play a crucial role in maintaining ecological balance. The application of unmanned aerial vehicle (UAV) imaging technology in wildlife monitoring has become increasingly mature. However, due to the influence of natural lighting conditions and the complex and changeable wild environment, it is difficult to obtain high-quality cervid images using single-spectrum imaging technology. Therefore, this paper proposes an image fusion algorithm based on the DenseFuse network. By utilizing the multispectral imaging equipment carried by unmanned aerial vehicles (UAVs), the algorithm fuses infrared images with visible light images while preserving the contour information of the infrared images and the appearance information of the visible light images, thereby improving the quality of monitoring images. Based on the wild cervid image dataset, this paper employs multiple image fusion strategies for experiments and conducts a detailed comparison of the fusion effects between infrared and visible light images. The experimental results show that the comprehensive evaluation index obtained by using the l_1 -norm fusion strategy is the best, and the average information entropy of the fused images reaches 6.965. This result indicates that the proposed UAV multi-source image fusion algorithm can provide reliable technical support for wildlife monitoring.

在全球生物多样性锐减的背景下,强化对野生动物资源的动态监测与系统性保护已成为维系生态安全与生物安全的战略举措^[1]。作为生态系统中关键物种的野生鹿科(Cervidae)动物,其种群稳定性正持续遭受气候变化、栖息地破碎及人类干扰等多重危机的叠加影响^[2-3]。相较于传统的人工巡查方式,无人机影像技术凭借其更大的覆盖范围和更高的监测通量,在野生动物监测领域的应用范围正逐步拓展^[4-5]。但是,受自然光照条件和野外环境复杂多变的影响,单一光谱成像技术很难得到高质量的野生鹿科动物图像。

目前,在动物监测中,无人机最为常用的监测手段是可见光成像技术。例如,通过无人机可见光图像对白额雁(*Anser albifrons*)^[6]、绵羊^[7]和牛^[8]的数量进行统计。结合人工智能技术,对无人机采集的可见光图像展开分析,从而实现非洲野生动物^[9]、恒河鳄(*Crocodylus palustris*)^[10]和沙漠动物^[11]的有效监

测。值得注意的是,在上述研究中,无人机大多在空旷区域开展动物数据的采集工作,很少出现动物被环境遮挡的情况。然而,在森林等植被茂盛的地区,使用无人机可见光成像进行野生鹿科动物监测时,由于树木遮挡,野生鹿科动物很难被完整监测到。相较于传统可见光成像技术,红外热成像技术依托红外辐射独特的物理穿透特性,在光照不足、植被冠层遮蔽等场景下仍能保持稳定的探测能力。这种热辐射感知机制突破了可见光成像的光学局限性,尤其在野生动物监测领域,当动物体色与背景环境形成光学伪装,或夜间可见光监测失效时,红外成像系统通过探测目标体表与环境的温差(通常哺乳动物体表温度比环境高 5~10 °C),仍能实现有效识别^[12]。Lyu *et al.*^[13]通过 Faster R-CNN 网络检测红外图像中的鹿,并针对红外图像分辨率低的问题,集成了小尺度锚框和多尺度特征图,提高了检测小物体

的准确性。但是,由于红外成像技术主要是依靠检测物体散发的热量,使得红外图像缺乏野生鹿科动物的纹理细节。相比之下,可见光可以提供鹿科动物高分辨率纹理细节信息。因此,针对野生鹿科动物监测场景中存在的多种成像技术互补性需求,亟需一种多光谱图像融合技术,将野生鹿科动物的可见光图像和热红外图像进行融合,既保证复杂林冠环境下动物主体的稳定识别,又能有效保留角部纹理、毛色斑纹等物种鉴别关键特征。

图像融合是将来自不同传感器获取的图像信息合并为单一图像^[14-15],可以提高图像质量、增强图像信息。因此,红外图像和可见光相融合后的融合图像能够吸收两者的优点,在保证分辨率的同时,也能保留清晰的轮廓特征和更多细节信息。Li *et al.*^[16]提出以一种基于深度学习的红外图像和可见光图像融合的算法,其中将源图像分为基础部分和高频部分并通过不同的融合策略以获得融合图像。Liu *et al.*^[17]通过自适应算法提高了红外图像对比度,但是可能会导致一些图像细节被压缩或丢失。为了让融合图像保留更多源图像的信息,杨莘等^[18]提出一种端到端的双融合路径生成对抗网络,其中将两幅源图像直接输入到网络的每一层,以提取更多的源图像特征信息。王昱婷等^[19]提出的融合模型 DAPR-Net 可以获得更清晰的目标细节和更明确的目标信息,并通过双注意力特征提取模块 AFEM,增强了在低光场景下的检测效果,但是在夜光场景下,可见光图像效果不佳。谢一博等^[20]通过双目异型成像系统和双尺度融合算法对红外与可见光图像进行融合,获得了信息丰富、质量更佳的融合图像。

目前,在野生动物的相关研究中,使用红外图像

与可见光图像相融合的研究有限。为减少野生动物检测中的遮挡问题,提高野生鹿科动物的监测效率,本研究提出一种基于 DenseFuse^[21] 图像融合网络,用于将无人机采集到的红外野生动物图像和可见光野生动物图像融合,拟解决检测过程中存在的遮挡和目标分辨率低造成检测效果不佳的问题。

1 实验数据采集与制作

1.1 视频采集

使用大疆经纬 M300 RTK 无人机,在保护区航拍采集 4 种鹿科动物的视频。2020 年 11 月,在吉林省延吉市的汪清自然保护区拍摄梅花鹿(*Cervus nippon*)和马鹿(*C. elaphus*)。2020 年 12 月,在内蒙古自治区根河市的汗马国家级自然保护区拍摄野生驯鹿(*Rangifer tarandus*)和狍(*Capreolus pygargus*)。无人机飞行高度越高,视频采集到的范围越大,能观察到的动物数量越多,但是飞行高度过高会导致图像中的动物太小,增加检测和识别的难度。通过不同飞行高度的图像效果对比,确定飞行高度为 30 ~ 200 m。

1.2 图像制式

拍摄大型哺乳动物时,变焦、广角和热成像相机同时录制,可同时采集 4 种视频,即以 SCRN(图 1(a))、THRM(图 1(b))、WIDE(图 1(c))和 ZOOM(图 1(d))为后缀的视频,其中 WIDE 用广角相机拍摄,ZOOM 用长焦相机拍摄,THRM 和 SCRN 为热成像相机拍摄的热红外视频,而 SCRN 图像是由 THRM 原始分辨率为 640 × 512 的视频填充黑边并拉伸至 1 920 × 1 080 分辨率视频生成。采集的红外视频中的部分图像见图 2。

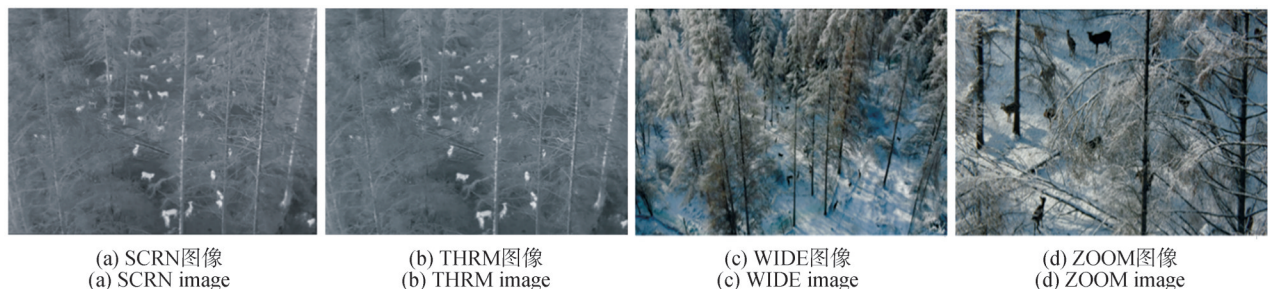


图1 无人机视频4种输出形式

Figure 1 Four output forms of drone video outputs

使用软件从视频中截取视频帧,截取频率为每 2 s 一帧,人工去除没有拍摄到动物或者成像效果差的图片,得到对应的 4 种制式图像。其中,THRM 和

SCRN 制式的图像使用热红外成像,获取到的动物轮廓更清晰,而 WIDE 和 ZOOM 制式图像保留了动物详细的外貌特征。

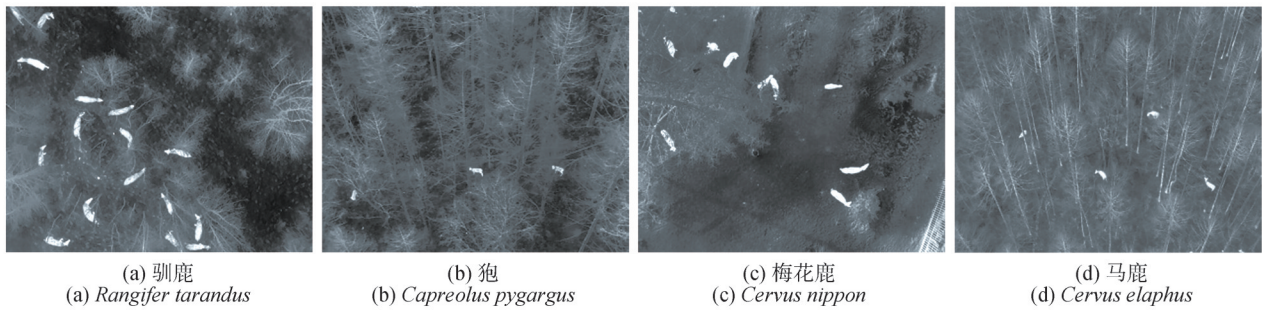


图2 4种鹿科动物红外图像示例

Figure 2 Examples of infrared images of four species of Cervidae animals

1.3 图像配准

使用图像融合算法对红外和可见光图像融合，需要选取2种制式的图像作为原始图像。经过对比，可见光图像ZOOM制式和红外图像THRM制式的视角更接近，故选择THRM和ZOOM制式图像进行融合。

如图3所示，融合要求2种图像尺寸一致且内容相互对齐，使用了特征匹配的配准算法将可见光图像对齐至红外图像，并对红外图像进行裁切后缩放至1920×1080分辨率。通过上述步骤，在4种动物图像中得到2200对配准的红外和可见光图像。

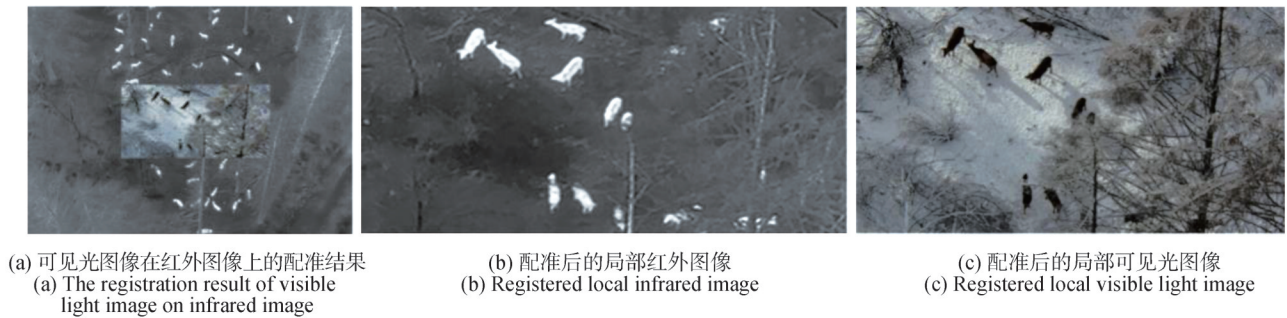


图3 图像配准以及内容对齐的红外和可见光图像

Figure 3 Image alignment as well as content-aligned infrared and visible images

1.4 数据制作

使用图像融合的方法对红外和可见光图像进行融合，且对应的红外和可见光图像中动物坐标保持一致。在得到的鹿科动物数据集中，红外图像组成如表1所示，其中马鹿498张，梅花鹿767张、狍481张、驯鹿454张。将2200张图片以4:1的比例随机划分成训练集和验证集，其中训练集1760张，验证集440张，数据集中的每张红外图像均有对应的可见光图像。

表1 数据集组成

Table 1 Composition of the data set

物种 Species	图像数量/张 Images number
马鹿 <i>Cervus elaphus</i>	498
梅花鹿 <i>Cervus nippon</i>	767
狍 <i>Capreolus pygargus</i>	481
驯鹿 <i>Rangifer tarandus</i>	454
总计 Total	2200

2 DenseFuse 融合网络

红外和可见光图像的融合图像需要从2种图像中提取出各自的特征，通过特定的融合方法将这些特征融合到新的图像之中。本研究使用DenseFuse深度学习网络来融合可见光和红外图像。

2.1 DenseFuse 融合原理

DenseFuse是一种基于CNN模块和DenseBlock的新型深度学习网络，与传统的深度学习网络相比，其使用了Encoder模块和Decoder模块，与网络中的卷积层、融合层相连接^[21]。虽然这种网络只能融合2种灰度图像，但是在融合红外和可见光图像时，我

们将可见图像的R、G、B三个通道分别视为3张灰度图像,与红外图像融合后,再作为R、G、B三个通道的图像输出,也能完成红外和可见光图像的融合。这种体系能够在红外和可见光图像中获取到更多有用的特征,并设计了2种融合策略,能够有效提高图像融合的性能。

2.2 DenseFuse网络结构

DenseFuse的网络架构包括3个部分:Encoder编码层、Fusion融合层和Decoder解码层。DenseFuse网络结构如图4所示。首先是编码层,包含2个部分,分别是C1层和DenseBlock层,它们用来特征提

取。C1使用了 3×3 的卷积核,DenseBlock由3层卷积层组成,卷积核尺寸也为 3×3 ,步长为1,因此得到的特征图尺寸不变,使用反射模式填充图像。每一个卷积层使用16个卷积核,故特征图为16通道。编码层有两点优势,一是因为卷积核尺寸为 3×3 ,步长和padding为1,使得输入和输出的尺寸一致,这样输入图像的大小无限制;二是DenseBlock模块可以在编码层尽可能多地保存深层特征,这些都有助于后面的融合策略。编码层采用4层卷积层,将融合层后的输出当作输入,来重建最终的融合图像。

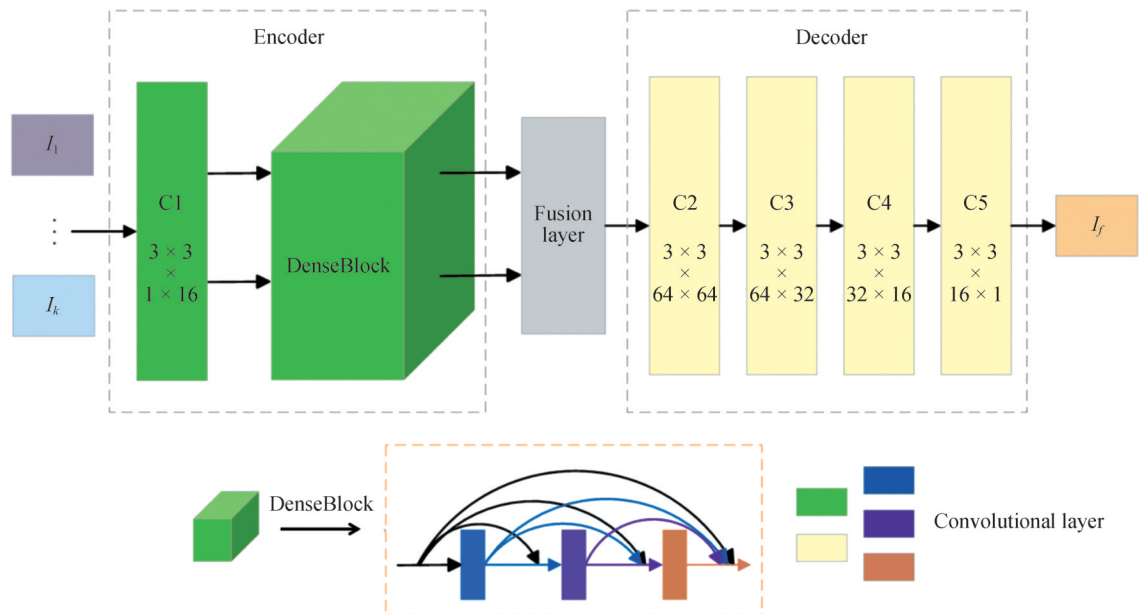


图4 DenseFuse网络结构

Figure 4 DenseFuse network structure

2.3 损失函数

在训练阶段,暂时忽略融合层,只考虑编码层和解码层,训练这两层用于重建输入图片。在编码和解码层的权重固定后,采用自适应特征融合策略进行融合。在训练时,使用损失函数(L)来进行反向传播,损失函数计算如下,

$$L = \lambda L_{ssim} + L_p \tag{1}$$

式中: L_{ssim} 为结构相似性损失,表示输入和输出差异性表达式, $L_{ssim} = 1 - SSIM(O, I)$,其中SSIM函数为相似性操作; L_p 为像素损失,是输入输出的欧式范数, $L_p = \|O - I\|_2$;L为 L_{ssim} 和 L_p 的加权组合,权重为 λ 。

2.4 融合策略

DenseFuse的核心是融合策略,编码层输出为高语义的特征图,融合策略决定了之后从解码层重建的图像与输入图像的相关程度,即融合性能。本研究使用addition和 l_1 -norm两种融合策略。

addition计算公式为

$$f^m(x, y) = \sum_{i=1}^k \phi_i^m(x, y) \tag{2}$$

式中: $m \in \{1, 2, \dots, M\}$, $M = 64$ 代表特征图的通道数; $k \geq 2$ 代表输入的索引,表示编码层从输入图像提取出来的特征图; f^m 表示融合后的特征映射。

addition的融合策略虽然有效,但是这样的操作

对于高语义特征融合效果不佳。故本研究使用了一种基于 l_1 -norm和softmax算法的新融合策略应用到融合网络中。

在新的融合策略中,首先计算活动等级图 C_i ,由 l_1 范数和基于块的平均算子计算,

$$C_i(x,y) \parallel \phi_i^{l_1}(x,y) \parallel_1. \quad (3)$$

通过基于块的平均算子计算能够得出最终活动水平图 \hat{C}_i 。 \hat{C}_i 计算如下,

$$\hat{C}_i(x,y) = \frac{\sum_{a=-r}^r \sum_{b=-r}^r C_i(x+a,y+b)}{(2r+1)^2}. \quad (4)$$

其中 r 决定块大小,本研究中 r 取1。

通过 \hat{C}_i 能得到融合后的特征图 f^m , f^m 计算如下,

$$f^m(x,y) = \sum_{i=1}^k \frac{\hat{C}_i(x,y)}{\sum_{n=1}^k \hat{C}_n(x,y)} \times \phi^m(x,y). \quad (5)$$

3 结果与分析

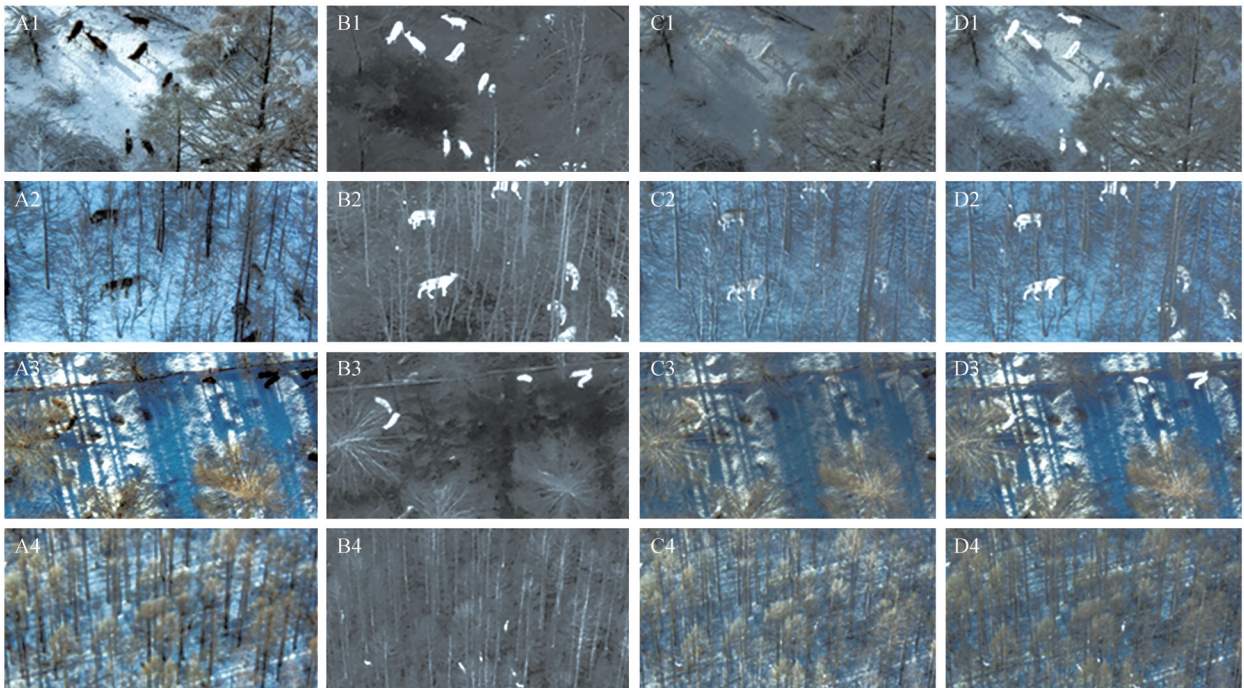
3.1 网络训练设置

采用迁移学习的方法,在COCO数据集上进行

预训练,训练集79 000张图像,测试集1 000张图像。随后,在鹿科动物数据集中随机抽取1 000张图像对网络进行训练。由于DenseFuse的总体损失是由SSIM损失和像素损失加权相加而得到,故需定义加权系数(λ),其中 λ 越大,表示SSIM损失在训练过程中权重越大。在本研究中,设置 $\lambda = 100$ 。

3.2 图像融合结果

将4种大型哺乳动物相互对应的THRM红外图像和ZOOM可见光图像进行融合。将红外图像和可见光图像输入进DenseFuse融合网络,分别使用2种特征融合策略,可以得到对应的融合图像(图5)。通过观察图像可知,可见光图像保留了动物的外貌花纹等细节信息,但是容易被树木遮蔽;红外图像保留了更多的轮廓信息,但是丢失了外貌特征。相比可见光图像和红外图像,使用DenseFuse融合网络得到的2种融合图像,不仅保留了外貌信息,还避免了被树木遮蔽,而且由于背景信息保留得更好,在后续的目标检测算法中能够更好地区分前景和背景,提高检测的准确性。相比于addition策略, l_1 -norm策略下融合图片的动物外貌更清晰,效果更佳。



A. 梅花鹿;B. 驯鹿;C. 狗;D. 马鹿;1. 可见光图像;2. 红外图像;3. addition融合图像;4. l_1 -norm融合图像。

A. *Cervus nippon*; B. *Rangifer tarandus*; C. *Capreolus pygargus*; D. *Cervus elaphus*; 1. Visible light image; 2. Infrared image; 3. addition fusion image; 4. l_1 -norm fusion image.

图5 4种鹿科动物图像DenseFuse融合前后对比

Figure 5 Comparison of four Cervidae animals images before and after DenseFuse fusion

3.3 融合图像质量评价

使用一些数据指标如 $SSIM$ ^[22-23]、信息熵^[24]、均方误差(MSE)和均方根误差(RMSE)和归一化均方根误差(NRMSE)等来量化融合图像的质量。在实验中增加2组对照组,分别为 DeepFuse 网络融合图

片和 VIFNet 融合图片。在使用指标进行计算时,需要将图像转换成灰度图像。如图6所示,每种动物各选取20组图像进行计算,每组包含6张图片。在计算完成后,记录所有图像的平均指标。

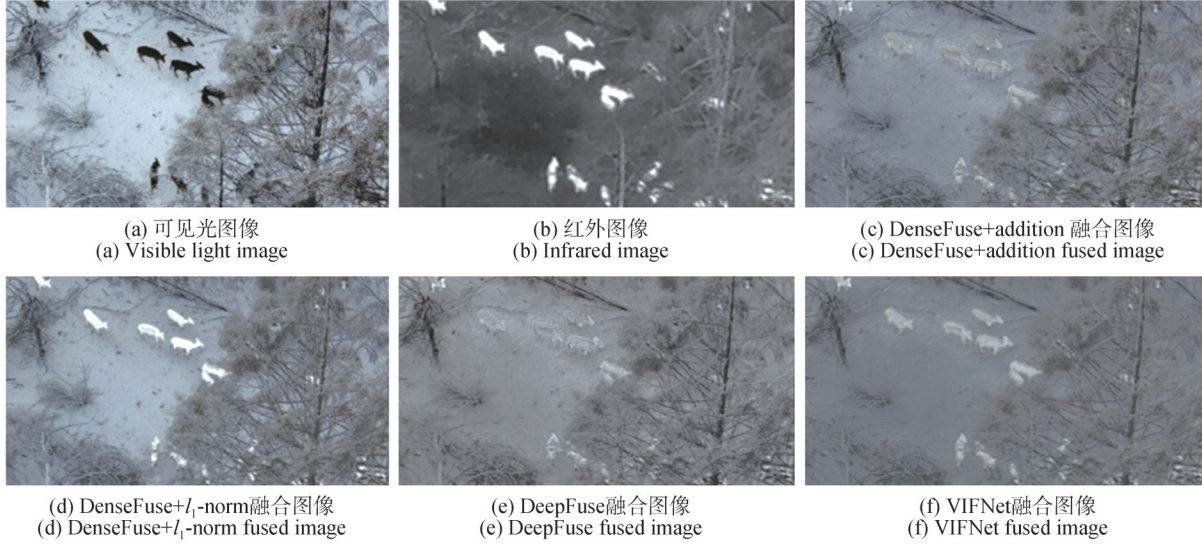


图6 可见光与红外图像融合示例

Figure 6 Example of visible light and infrared image fusion

$SSIM$ 表示两张图片的相似性,该指标通过3个参数来衡量图像的相似度,设两张图片分别为 x 和 y ,分别得到亮度相似性 $l(x,y)$ 、对比度相似性 $c(x,y)$ 和结构比较相似性 $s(x,y)$,最终得到 $SSIM$ 指标。其中,亮度相似性、对比度相似性、结构比较相似性和 $SSIM$ 的计算式为

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}, \quad (6)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}, \quad (7)$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}, \quad (8)$$

$$SSIM(x,y) = [l(x,y)]^\alpha [c(x,y)]^\beta [s(x,y)]^\gamma. \quad (9)$$

式中: μ_x 表示 x 的平均值, μ_y 表示 y 的平均值, σ_x^2 和 σ_y^2 分别表示 x 和 y 的方差。 $SSIM$ 取值范围为 $[-1, +1]$,值越大,表示相似度越高,图像融合质量越好。一般选取 $\alpha = \beta = \gamma = 1$ 、 $c_3 = c_2/2$ 以简化 $SSIM$ 计算,简化后的 $SSIM$ 计算式为

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}. \quad (10)$$

计算 addition 策略的融合图像、 l_1 -norm 策略的融

合图像、DeepFuse 融合图像和 VIFNet 融合图像分别与红外和可见光图像的 $SSIM$,结果如图7所示。

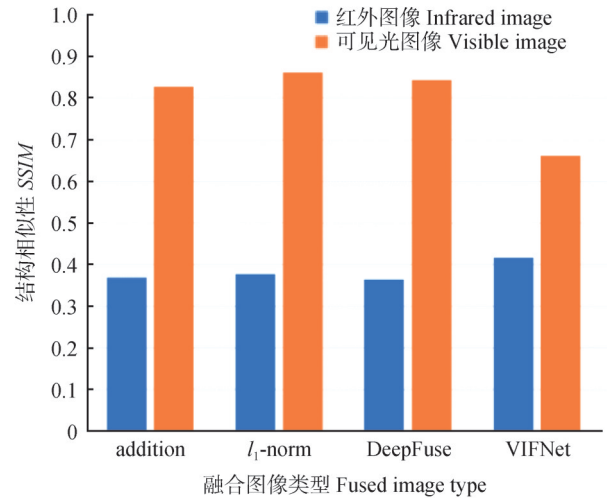


图7 4种融合图像与红外和可见光图像的 $SSIM$

Figure 7 $SSIM$ calculation results of the four fused images with the infrared images and visible light images

由图7可知,相比于 addition 策略的融合图像、DeepFuse 融合图像和 VIFNet 融合图像, l_1 -norm 策略的融合图像对于融合后的细节保留能力最佳。

MSE、RMSE 和 NRMSE 是常用的图像融合质量

评估指标,用于衡量融合图像与原始图像之间的差异。这些指标基于像素误差计算,数值越小,表示融合图像与原始图像的相似度越高,融合效果越好。其中,MSE的计算公式为

$$E = \frac{1}{N} \sum_i^N (x_i - y_i)^2. \quad (11)$$

式中: x_i 和 y_i 是两张图像的像素点。

表2 融合图像与红外及可见光图像的MSE、RMSE和NRMSE评估结果

Table 2 Comparison of MSE, RMSE and NRMSE between fusion images and infrared and visible light images respectively

图像融合方法 Image fusion method	红外图像评估 Infrared image evaluation			可见光图像评估 Visible image evaluation		
	MSE	RMSE	NRMSE	MSE	RMSE	NRMSE
addition	1 824	42.7	0.55	1 423	37.7	0.29
l_1 -norm	2 253	47.5	0.62	1 256	35.4	0.28
DeepFuse	2 916	54.0	0.69	1 322	36.5	0.28
VIFNet	2 270	47.6	0.62	2 332	48.3	0.37

信息熵(entropy)表示图像灰度级集合的比特平均数,描述了图像包含信息的多少。信息熵越大,表示包含的信息越多。信息熵计算公式为

$$H(A) = -\sum_a P_A(a) \log P_A(a). \quad (12)$$

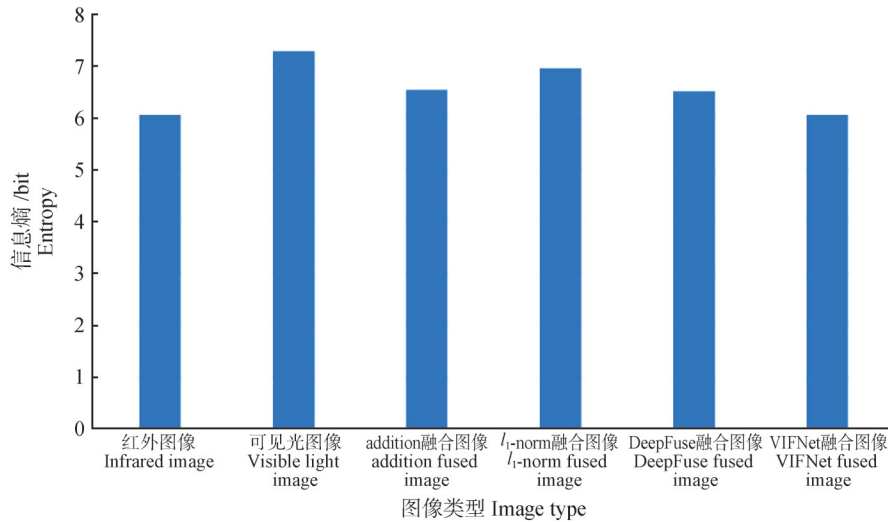


图8 6种图像的平均信息熵

Figure 8 Average information entropy of six images

经过以上几个指标的对比, l_1 -norm融合图像参数最优,在融合了红外和可见光图像后,携带的信息最多。结合对样本的观察, l_1 -norm融合图像对动物轮廓信息保留得最好,能够有效地减小树木遮蔽,提高后续目标检测精度。

4 结论

本研究提出了一种基于DenseFuse网络的图像

融合算法,使用 l_1 -norm融合策略对红外图像和可见光图像进行融合。DenseFuse网络生成的融合图像结合了红外图像和可见光图像的优势,生成了适用于野外有遮蔽场景的融合图像数据。这些融合图像数据能够有效地提升动物检测性能,解决无人机拍摄时动物容易被树木冠层遮挡以及动物毛色与环境相近无法区分的问题。使用DenseFuse深度学习网络对2种图像融合后,得到包含2种图像特征的融合图像。

RMSE只是将MSE进行平方根计算,而NRMSE是将RMSE的值映射到0~1。MSE、RMSE和NRMSE对比结果如表2所示。由表2可知,在红外图像中,addition融合图像、 l_1 -norm融合图像、VIFNet融合图像和DeepFuse融合图像后的差异度依次增高;在可见光图像中, l_1 -norm融合图像的差异最小,可以保留更多的鹿科动物外貌信息。

将红外图像、可见光图像、addition融合图像、 l_1 -norm融合图像、VIFNet融合图像和DeepFuse融合图像分别进行信息熵计算,计算结果如图8所示。由图8可知,可见光图像的信息熵最大;在融合图像中, l_1 -norm融合图像的信息熵最大,包含的信息最多。

融合算法,使用 l_1 -norm融合策略对红外图像和可见光图像进行融合。DenseFuse网络生成的融合图像结合了红外图像和可见光图像的优势,生成了适用于野外有遮蔽场景的融合图像数据。这些融合图像数据能够有效地提升动物检测性能,解决无人机拍摄时动物容易被树木冠层遮挡以及动物毛色与环境相近无法区分的问题。使用DenseFuse深度学习网络对2种图像融合后,得到包含2种图像特征的融合图像。

经过对比实验, l_1 -norm融合图像相比于addition、DeepFuse和VIFNet融合图像,各项指标参数更优,并由信息熵的参数可知, l_1 -norm融合图像在轮廓信息接近红外图像的同时,其信息熵更接近可见光图像。

参考文献:

- [1] DU N, FATHOLLAHI-FARD A M, WONG K Y. Wildlife resource conservation and utilization for achieving sustainable development in China: main barriers and problem identification [J]. *Environmental Science and Pollution Research*, 2023. DOI: 10.1007/s11356-023-26982-7.
- [2] LINNELL J D C, CRETOIS B, NILSEN E B, *et al.* The challenges and opportunities of coexisting with wild ungulates in the human-dominated landscapes of Europe's Anthropocene [J]. *Biological Conservation*, 2020, 244: 108500.
- [3] FORSYTH D M, COMTE S, DAVIS N E, *et al.* Methodology matters when estimating deer abundance: a global systematic review and recommendations for improvements [J]. *The Journal of Wildlife Management*, 2022, 86(4): e22207.
- [4] DE KOCK M E, POHŮNEK V, HEJCMANOVÁ P. Semi-automated detection of ungulates using UAV imagery and reflective spectrometry [J]. *Journal of Environmental Management*, 2022, 320: 115807.
- [5] LI X H, HUANG H L, SAVKIN A V. Autonomous navigation of an aerial drone to observe a group of wild animals with reduced visual disturbance [J]. *IEEE Systems Journal*, 2022, 16 (2) : 3339-3348.
- [6] OGAWA K, LIN Y T, TAKEDA H, *et al.* Automated counting wild birds on UAV image using deep learning [C]//2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, July 11-16, 2021. Brussels: IEEE, 2021: 5259-5262.
- [7] SARWAR F, GRIFFIN A, REHMAN S U, *et al.* Detecting sheep in UAV images [J]. *Computers and Electronics in Agriculture*, 2021, 187: 106219.
- [8] DE LIMA WEBER F, DE MORAES WEBER V A, DE MORAES P H, *et al.* Counting cattle in UAV images using convolutional neural network [J]. *Remote Sensing Applications: Society and Environment*, 2023, 29: 100900.
- [9] PETSOT, JAMISOLA R S, MPOELEN D, *et al.* Individual animal and herd identification using custom YOLO v3 and v4 with images taken from a UAV camera at different altitudes [C]//2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP), October 22-24, 2021. Nanjing: IEEE, 2021: 33-39.
- [10] DESAI B, PATEL A, PATEL V, *et al.* Identification of free-ranging mugger crocodiles by applying deep learning methods on UAV imagery [J]. *Ecological Informatics*, 2022, 72: 101874.
- [11] CHEN C R, EDIRISINGHE E A, LEONCE A, *et al.* Deep neural networks based multiclass animal detection and classification in drone imagery [C]//2023 International Symposium on Networks, Computers and Communications (ISNCC), October 23-26, 2023. Doha: IEEE, 2023: 1-8.
- [12] LI W X, CHEN Q, GU G H, *et al.* Visible-infrared image matching based on parameter-free attention mechanism and target-aware graph attention mechanism [J]. *Expert Systems with Applications*, 2024, 238: 122038.
- [13] LYU H T, QIU F, AN L, *et al.* Deer survey from drone thermal imagery using enhanced faster R-CNN based on ResNets and FPN [J]. *Ecological Informatics*, 2024, 79: 102383.
- [14] MA J Y, MA Y, LI C. Infrared and visible image fusion methods and applications: a survey [J]. *Information Fusion*, 2019, 45: 153-178.
- [15] JIN X, JIANG Q, YAO S W, *et al.* A survey of infrared and visual image fusion methods [J]. *Infrared Physics & Technology*, 2017, 85: 478-501.
- [16] LI H, WU X J, KITTLER J. Infrared and visible image fusion using a deep learning framework [C]//2018 24th International Conference on Pattern Recognition (ICPR), August 20-24, 2018. Beijing: IEEE, 2018: 2705-2710.
- [17] LIU F, GUAN S, YU K K, *et al.* Infrared target detection based on the fusion of mask R-CNN and image enhancement network [C]//2022 China Automation Congress (CAC), November 25-27, 2022. Xiamen: IEEE, 2022: 2011-2016.
- [18] 杨莘, 田立凡, 梁佳明, 等. 改进双路径生成对抗网络的红外与可见光图像融合 [J]. *电子与信息学报*, 2023, 45(8): 3012-3021.
YANG S, TIAN L F, LIANG J M, *et al.* Infrared and visible image fusion based on improved dual path generation adversarial network [J]. *Journal of Electronics & Information Technology*, 2023, 45(8): 3012-3021.
- [19] 王昱婷, 刘志明, 万亚平, 等. 基于可见光与红外图像的弱光条件下目标检测 [J]. *计算机工程*, 2024, 50(8): 270-281.
WANG Y T, LIU Z M, WAN Y P, *et al.* Target detection under low light conditions based on visible and infrared images [J]. *Computer Engineering*, 2024, 50(8): 270-281.
- [20] 谢一博, 程进, 周顺, 等. 夜视环境下红外与可见光图像真彩色快速融合方法研究 [J]. *激光与红外*, 2024, 54(1): 136-147.
XIE Y B, CHENG J, ZHOU S, *et al.* Research on the fast fusion algorithm of true colour of infrared and visible images under night vision environment [J]. *Laser & Infrared*, 2024, 54 (1): 136-147.
- [21] LI H, WU X J. DenseFuse: a fusion approach to infrared and visible images [J]. *IEEE Transactions on Image Processing*, 2019, 28(5): 2614-2623.
- [22] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [23] ZHAO H, GALLO O, FROSIO I, *et al.* Loss functions for image restoration with neural networks [J]. *IEEE Transactions on Computational Imaging*, 2017, 3(1): 47-57.
- [24] CHEN X Q, ZHANG Q Y, LIN M H, *et al.* No-reference color image quality assessment: from entropy to perceptual quality [J]. *EURASIP Journal on Image and Video Processing*, 2019, 2019(1): 77.