

面向天基红外预警的高动态弱小目标 LSTM 检测方法研究

翟光, 胡圣冉, 孙一勇*

(北京理工大学 宇航学院, 北京 100081)

摘要:在天基红外预警任务中,高动态弱小目标具有成像尺寸小、运动规律未知的特点。现有红外弱小目标探测任务主要关注匀速直线运动目标检测问题,对高动态目标的有效检测算法尚需进一步开发。针对天基红外预警任务中高动态非线性运动目标检测问题,提出了一种基于 LSTM 的红外弱小目标检测算法。首先设计了提取可疑目标位置信息的预处理方法,解决了 LSTM 网络结构与序列图像不匹配的问题;然后,针对传统算法难以检测非线性运动轨迹的问题,利用 LSTM 提取目标运动特征,实现序列图像中高动态目标的检测。通过与序列假设检验等算法的对比,在自研的红外序列图像数据集上验证了所提出的算法能够实现不低于 0.9347 的精确率与不低于 0.8633 的召回率。

关键词:天基预警; 目标检测; LSTM; 高动态

中图分类号: TP391 **文献标志码:** A **DOI:** 10.3788/IRLA20230010

0 引言

相比雷达探测,红外探测技术具有观测范围大、扫描周期短、抗干扰能力强的优势,在天基预警任务中具有重要意义。任务环境中目标与传感器距离一般较远,能量在传播时衰减显著,因此目标在红外系统中一般成像为弱小目标;同时,背景杂波及噪声干扰进一步降低了红外图像的信杂比 (Signal to Clutter Ratio, SCR) 和信噪比 (Signal to Noise Ratio, SNR),提升了目标检测的难度。近年来,复杂机动航天器技术快速发展,为红外弱小目标检测任务带来了新的挑战。

红外弱小目标检测旨在根据红外图像信息计算目标的位置,为目标跟踪等任务提供必要条件。根据单次检测所利用的图像数量,传统红外弱小目标检测算法可分为基于单帧图像的跟踪前检测算法 (Detect before Track, DBT) 与基于序列图像的检测前跟踪算法 (Track before Detect, TBD)。跟踪前检测算法一般先对单幅图像进行滤波处理,通过阈值分割方法提取疑似目标,在对当前图像中目标进行判别后进行跟踪等后续任务。Deshpande 等人^[1]设计了最大中值滤波 (Max-Medium Filtering, MMF) 抑制背景中的边界部

分;Chen 等人^[2]提出局部对比度测量 (Local Contrast Measure, LCM) 方法,注重于窗口内灰度值差异并以此为依据进行检测,Chen 等人^[3]、鲁晓峰等人^[4]基于 LCM 机制在噪声抑制及算法实时性等方面进行改进;跟踪前检测算法在检测时仅利用单帧图像信息,实时性较好,但仅适用于信噪比较高、背景较为简单的检测任务。检测前跟踪算法通常以假设的目标运动规律为先验知识,对序列图像中形成的可疑轨迹进行判断。Falconer^[5]使用直线检测中常用的 Hough 算法检测目标运动轨迹;Barniv^[6]根据贝叶斯理论,将动态规划方法应用于轨迹检测;Blostein 等人^[7]提出序列假设检验 (Sequential Hypothesis Testing, SHT) 算法,对可疑轨迹进行逐层剪枝;Zhang 等人^[8]、Pang 等人^[9]、Uzair 等人^[10]将图像空间域中的目标检测方法扩展到时空域以处理时序图像;检测前跟踪算法利用多帧图像中目标运动信息进行检测,可适用于低信噪比下的检测任务,但由于其常基于匀速直线运动假设进行判断,难以对复杂机动目标进行较高精度的检测。

伴随深度学习技术的发展,以 R-CNN 系列与 YOLO 系列^[11]为代表的目标检测算法实现了良好的效果,但其主要应用于形状、纹理等特征丰富的常规

收稿日期:2023-01-05; 修订日期:2023-04-30

作者简介:翟光,男,教授,博士,主要从事空间目标相对导航与控制方面的研究。

通讯作者:孙一勇,男,助理教授,博士,主要从事人工智能技术与应用方面的研究。

目标,对于红外弱小目标检测问题能力不足,为解决这一问题,Hou 等人^[12]、Dai 等人^[13]、蒋昕昊等人^[14]基于卷积神经网络 (Convolutional Neural Network, CNN) 及 YOLO 框架,在实时性方面进行改进以检测弱小目标。部分学者尝试利用长短期记忆网络 (Long Short-time Memory, LSTM) 提取时序特征的能力对目标轨迹进行检测,通常结合卷积神经网络提取目标空间特征; Baird 等人^[15]使用 CNN-LSTM 对小目标识别进行增强, CNN 用于提取空间信息, LSTM 负责分析空间信息的时间变化特性; Liu 等人^[16]借助 Conv-LSTM 算法,将 LSTM 的输入矢量扩展为二维矩阵以适应图像处理,同时采用 Bi-LSTM 结构以充分利用目标的时序信息。深度学习方法虽然能适应复杂机动情况,但其处理红外弱小目标检测任务存在诸多困难;深度学习方法对数据集规模要求较高,而红外弱小目标仍不存在公认的大规模数据集;深度学习方法常以 CNN 为基础,检测纹理、形状信息缺失的小目标能力不足;红外弱小目标检测任务往往需要较高的实时性,限制了深度学习算法的能力。因此,当前以深度学习方法处理红外弱小目标检测任务的研究相对较少。

针对现有方法中存在的上述问题,文中面向复杂机动的红外弱小目标检测任务,设计了一种基于 LSTM 的红外弱小目标检测 (LSTM-based Infrared Small-and-Dim Target Detection, LISTD) 算法,以中值滤波算法提取图像中的可疑点位置坐标,将其以矢量形式输入 LSTM 网络,从而在时序信息中提取目标轨迹的运动特征并完成目标检测。在自主设计制作的红外仿真数据集进行测试,验证算法的可行性和实用性。该算法具有如下特点:

1) 针对目标信号占比小的特点,设计预处理算法在二维红外图像矩阵中提取目标位置信息并转化为一维的矢量信息,使 LSTM 网络能够维持其较简单的结构并参与到图像处理中,相比扩展 LSTM 维度的方法有效降低计算量并减少计算耗时。

2) 以可疑目标坐标序列为输入,利用目标运动的时序特性,使用 LSTM 对序列图像中的运动目标进行识别和定位,并结合预处理结果修正 LSTM 输出的位置坐标,提升检测精度。根据高动态目标特性设计多种场景的序列图像训练集,提升训练模型的检测能力。

1 天基预警任务红外图像特征模型

天基红外预警卫星利用红外成像系统接收视场内的物体辐射的能量,经热电转化形成电信号,最后呈现为红外图像,图像中灰度值大小表征了不同物体之间的辐射温度的差异。包含目标的红外图像可建模,如公式 (1) 所示:

$$x(m, n) = x_T(m, n) + x_B(m, n) + x_N(m, n) \quad (1)$$

式中: $x(m, n)$ 为红外图像坐标 (m, n) 处的灰度值; $x_T(m, n)$ 、 $x_B(m, n)$ 、 $x_N(m, n)$ 分别为图像坐标 (m, n) 处因目标、背景、噪声信号所形成的灰度值。

单张红外图像包括以下三种信号:

1) 红外弱小目标信号。目标在成像平面上成像尺寸一般小于 9×9 , 在整张图像上面积占比极小,难以利用其形状、结构、纹理等信息,但其与背景间不存在空间相关性而在灰度上具有局部对比度。可使用高斯模型描述其信号,如公式 (2) 所示:

$$x_T(m, n) = s_E \exp \left\{ -\frac{1}{2} \left[\frac{(m - x_T)^2}{\sigma_x^2} + \frac{(n - y_T)^2}{\sigma_y^2} \right] \right\} \quad (2)$$

式中: $x_T(m, n)$ 为中心坐标为 (x_T, y_T) 的目标在坐标 (i, j) 处的灰度值; s_E 为目标信号幅值。即目标灰度值分布呈现出中心高,周围逐渐衰弱的特征。

2) 红外背景信号。背景在天基预警任务中体现为恒星、地表、云层,相比目标信号,背景在图像中占比较大,空间上变化较为平缓,是图像中的低频部分,但云层等背景的边界及角点易被误识别为目标。在较长时间的序列图像中,背景由于自身运动或预警卫星相机的扫视,会出现整体缓慢运动的现象。

信杂比是体现目标与背景杂波强度差异的指标,文中采用的信杂比计算方法如公式 (3) 所示:

$$SCR = \frac{|\mu_t - \mu_b|}{\sigma_b} \quad (3)$$

式中: μ_t 为目标范围内的灰度均值; μ_b 为背景图像的灰度均值, σ_b 为背景图像的灰度标准差。

3) 噪声信号。噪声主要来自视场内背景和探测系统本身,可细分为背景噪声、热噪声、载流子噪声等,大部分噪声近似满足加性的高斯噪声模型,其概率密度函数为:

$$p(x) = \frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(x-\bar{x})^2}{2\delta^2}} \quad (4)$$

式中: x 为加性噪声的灰度值; \bar{x} 为噪声平均幅值; δ 为噪声标准差。信噪比是体现目标与噪声强度的指标, 对于尺寸为 $M \times N$ 的图像, 文中采用的信噪比计算方法如公式 (5) 所示:

$$SNR = 10 \lg \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [g(i, j)]^2}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [g(i, j) - f(i, j)]^2} \quad (5)$$

式中: $g(i, j)$ 、 $f(i, j)$ 分别为噪声图像和原图像在坐标 (i, j) 处的灰度值。

显然, 噪声会在图像中形成一定数量的高灰度值点, 对弱小目标检测形成干扰。在序列图像中, 噪声体现出非连续和无规律的特点。

在红外弱小目标检测算法中, 目标中心幅值相对周围信号较高的特征是最为广泛使用的判断依据, 但基于该特性的相关检测算法存在目标信号弱小, 背景

边界、背景角点及噪声干扰明显等难点, 易造成较高的虚警率和较低的检测效率。因此, 有必要在检测中利用目标的运动特征。

在序列图像中, 背景的协同低速运动及噪声无规律出现的特点, 使连续运动的目标有可能被检出。对于复杂机动目标, 在没有关于目标运动规律的先验知识的情况下, 传统算法难以仅利用目标运动的连续性进行有效检测, 而以 LSTM 为代表的学习类算法具有处理该类问题的潜力。

2 算法介绍

2.1 LSTM 网络结构

基于多帧图像的传统检测前跟踪算法能够利用目标运动特征, 但多以目标匀速直线运动为假设, 难以处理复杂机动目标。为解决这一问题, 可引入 LSTM 网络, 利用其处理序列数据的能力提取目标运动特征, 进行目标检测。LSTM 网络单元结构如图 1 所示。

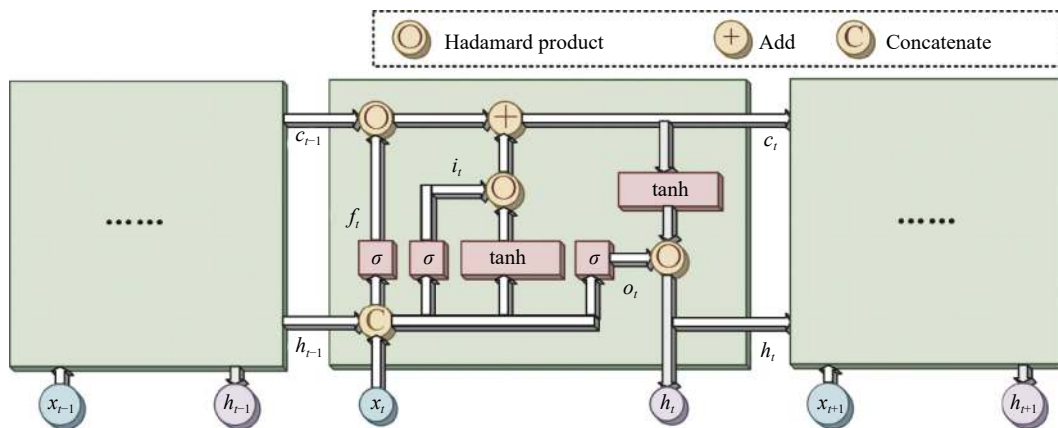


图 1 LSTM 网络单元结构

Fig.1 Structure of LSTM network unit

图 1 中, i_t 、 f_t 、 o_t 分别被称为 t 时刻单元的输入门、遗忘门和输出门; c_t 、 h_t 分别被称为 t 时刻的单元状态和隐藏状态; x_t 为 t 时刻单元的输入值。输入门、遗忘门和输出门的计算如公式 (6) 所示:

$$\begin{aligned} i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \end{aligned} \quad (6)$$

式中: W_{xi} 、 W_{xf} 、 W_{xo} 分别为计算输入门、遗忘门和输出门时输入值对应的权重矩阵; W_{hi} 、 W_{hf} 、 W_{ho} 分别为

计算输入门、遗忘门和输出门时隐藏状态对应的权重矩阵; b_i 、 b_f 、 b_o 分别为计算输入门、遗忘门和输出门时采用的偏置量; σ 和 \tanh 为激活函数, 对于范围为 $(-\infty, +\infty)$ 的实数, 前者将其转化至 $[0, 1]$, 后者则转化为 $[-1, 1]$ 。

单元状态和隐藏状态的计算如公式 (7) 所示:

$$\begin{aligned} c_t &= f_t \circ c_{t-1} + i_t \circ \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ h_t &= o_t \circ \tanh(c_t) \end{aligned} \quad (7)$$

式中: W_{xc} 、 W_{hc} 为计算单元状态时输入值和隐藏状态对应的权重矩阵; b_c 为计算单元状态时采用的偏置

量; \circ 为哈达玛积。

在高动态弱小目标检测任务中, 序列图像上的目标运动可能呈现非线性特征, 将目标轨迹假设为直线的检测方法不能有效检测高动态目标; 由于非合作目标运动规律未知, 无法建立描述其复杂运动特征的准确模型, 因此不能有效地利用时域信息检测目标。考虑 LSTM 在处理时序数据问题中具有良好表现, 对高动态弱小目标检测具有一定的应用价值。但 LSTM 输入一般为矢量数据, 无法直接处理涉及二维矩阵的

图像相关问题。目前的处理思路往往为使用 CNN 对图像信息进行提取并降维为矢量或对 LSTM 输入进行维度扩充, 但这些方法都涉及对二维图像的深度处理, 运算复杂性较高, 耗时较多, 且由于弱小目标空间域特征不明显, 这类方法对弱小目标检测问题往往不能发挥最佳作用。

针对上述问题, 文中提出 LISTD 算法, 根据对序列图像中可能存在的目标进行检测, 输出目标的位置, 其总体流程如图 2 所示。

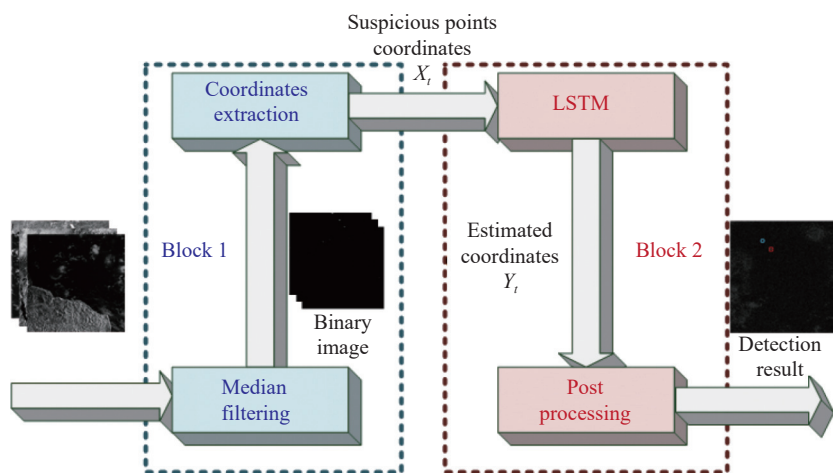


图 2 LISTD 算法流程

Fig.2 Process of LISTD algorithm

文中提出的 LISTD 算法具有两个模块: Block 1 为预处理模块, 采用中值滤波算法提取原始图像中可疑目标的位置信息, 将图像信号转化为矢量信息, 便于后续的运动特征提取; Block 2 为轨迹提取模块, 是算法的核心部分, 使用 LSTM 对目标运动特征进行学习与提取, 并估计目标的存在性及位置坐标, 最后输出检测结果。

2.2 基于中值滤波的预处理算法

针对图像信号不能直接用于经典 LSTM 网络的问题, 考虑弱小目标检测任务中目标占图像面积小的特点, 设计了基于中值滤波的预处理算法, 将二维的图像信号提取为一维的矢量信息作为 LSTM 网络的输入。预处理模块如图 3 所示。

为了实现快速提取可疑目标的位置信息, 采用如图 3(b)所示的中值滤波方法。图 3(b)中标注了滤波核各单元对应位置的原图像灰度值, 若原始图像在图

像坐标 (m, n) 处的灰度值记为 $x(m, n)$, 平滑后的图像灰度值记为 $y(m, n)$, 则滤波操作如公式 (8) 所示:

$$y(m, n) = \text{median}\{x(m-2, n-2), x(m-2, n-1), \dots, x(m+2, n+2)\} \quad (8)$$

然后将 $y(m, n)$ 与 $x(m, n)$ 做差分, 如公式 (9) 所示:

$$z(m, n) = x(m, n) - y(m, n) \quad (9)$$

式中: $z(m, n)$ 为差分后图像在 (m, n) 处的灰度值, 对其进行二值化, 如公式 (10) 所示:

$$t(m, n) = \begin{cases} 1, & z(m, n) > T \\ 0, & z(m, n) < T \end{cases} \quad (10)$$

式中: $t(m, n)$ 为二值化图像在 (m, n) 处的灰度值; T 为二值化阈值, 其计算方法为:

$$T = \mu + k\sigma \quad (11)$$

式中: μ 、 σ 分别为原始图像灰度的均值和标准差; k 为经验参数, 一般取值 10~15。为充分发挥预处理算法的作用, 避免提取信息过少导致无法进行运算或提取

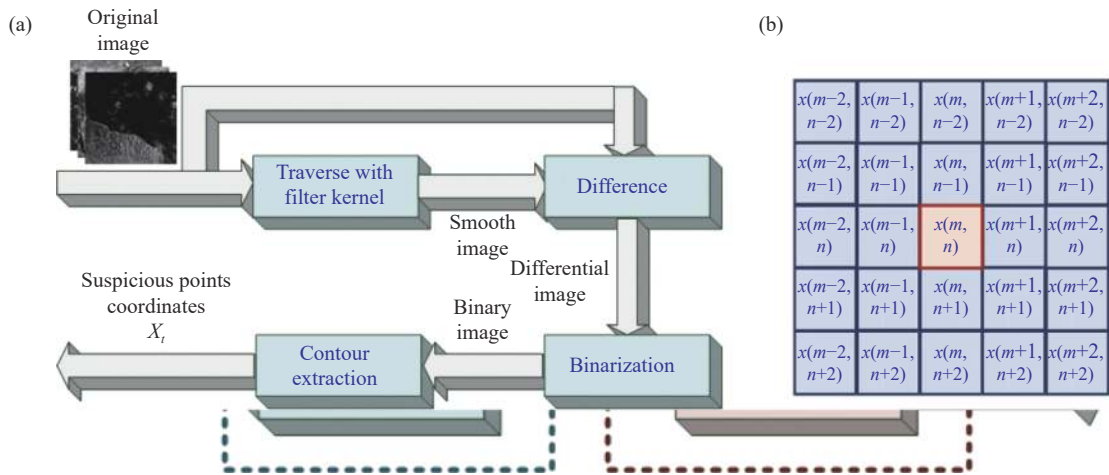


图 3 预处理模块。(a)模块流程; (b)滤波核

Fig.3 Preprocessing block. (a) Block process; (b) Filter core

信息过多干扰运算,文中设计了针对每帧图像的 k 值自适应机制,即以 $k = 12$ 为初值,在循环中以1为步长调整 k 值,直至二值化图像中灰度值为1的区域数量为 $n_{\text{point,min}} \sim n_{\text{point,max}}$ 个,或达到预设的循环次数。 $n_{\text{point,min}}$ 、 $n_{\text{point,max}}$ 为人为设定的最小点数量和最大点数量,文中取 $n_{\text{point,min}} = 10, n_{\text{point,max}} = 20$ 。

对于二值化图像中各灰度值为1的区域,计算其质心,作为可疑目标点的位置坐标。对于当前帧中的每个可疑目标点,以 r 为半径在前一帧提取到的可疑目标点坐标中搜索,取最近点计算其距离,记为该点的猜想速度值 V_g ;再取当前帧可疑点中 V_g 最大的 n_{point} 个点坐标列为矢量 $x_t = [x_1, y_1, x_2, y_2, \dots, x_{n_{\text{point}}}, y_{n_{\text{point}}}]$ 作为 LSTM 的输入,其中 x_i, y_i 为按 V_g 排序第 i 的可疑点的 x, y 坐标值;若二值化后可疑目标数目不足 n_{point} ,则以缺省值 d 填充补齐。前述 r, n_{point} 与 d 均为人为设定值,文中取 $r = 8, n_{\text{point}} = 8, d = -16$ 。

在预处理模块中,为适应 LSTM 网络结构,利用原始图像中目标点灰度值高于背景的特征,提取可疑目标位置信息,将二维的矩阵信号转化为一维的矢量信息;根据估计速度对可疑目标进行取舍,提高目标相关信息被输入至 LSTM 的概率;针对目标信号尺寸小的特点,在预处理阶段利用目标形状、尺寸等信息并在预处理后舍弃,仅保留目标的位置信息,将单帧图像所包含的大量信息转化固定长度的矢量,减轻了 LSTM 的运算负担,并突出了目标的运动特征。

2.3 基于 LSTM 的轨迹检测

借助 LSTM 提取时序特征的能力,使 LSTM 学习

目标轨迹的运动规律,并以此为判断依据,从随机出现的噪声点轨迹及协同低速运动的背景角点轨迹中区分真实的复杂运动目标点轨迹,最终对目标的存在进行判断,输出当前目标点坐标的估计值。

LSTM 输出的仅为估计目标点,对于图像中未出现目标点的情况,则以输出一个负的缺省值 d 来表示。不采取增加置信度项的方式来表示是否检测出图像中目标,主要出于三方面考虑:首先,缺省值方法相比置信度方法减少了 LSTM 网络内权重矩阵的规模,降低运算复杂度,提升算法的实时性;其次,在计算损失函数时,置信度的设计会带来较大影响,干扰 LSTM 的学习效率;最后,文中算法主要针对单一类型目标的检测,不涉及目标分类问题,无需设计置信度。LSTM 轨迹检测模块算法流程如图 4 所示。

图 4 中,将 t 时刻图像经预处理算法计算所得长度为 $2n_{\text{point}}$ 的可疑目标点坐标矢量 $x_t = [x_1, y_1, x_2, y_2, \dots, x_{n_{\text{point}}}, y_{n_{\text{point}}}]$ 作为 t 时刻 LSTM 的网络输入,以 128 为时间步长估计 t 时刻目标点坐标 $y_t = [x_{\text{pre}}, y_{\text{pre}}]$, x_{pre} 和 y_{pre} 即为所估计的目标点在图像平面中的 x, y 坐标值。

$y_t = [x_{\text{pre}}, y_{\text{pre}}]$ 作为 LSTM 根据时序特征对目标点坐标的估计值,相对于该时刻目标点真实坐标仍有一定误差,而预处理算法计算结果提供了较为准确的可疑目标点的位置坐标。为提升检测精度,对 LSTM 运算结果进行后处理,从预处理结果中选取最为接近的目标点坐标作为最终结果。其计算方法如公式(12)所示:

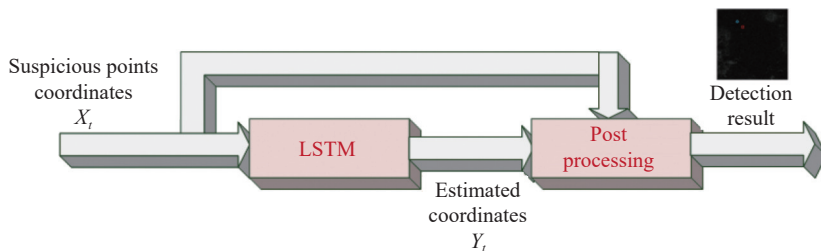


图 4 LSTM 检测模块

Fig.4 LSTM detection block

$$z_i = [x_{out}, y_{out}] = \begin{cases} [d, d], \min [(x_i - x_{pre})^2 + (y_i - y_{pre})^2]^{\frac{1}{2}} > T_d, \\ \quad \text{or}, x_{pre} < 0, \text{or}, y_{pre} < 0 \\ \arg \min_{(x_i, y_i)} [(x_i - x_{pre})^2 + (y_i - y_{pre})^2]^{\frac{1}{2}}, \text{else} \end{cases} \quad (12)$$

式中: $z_i = [x_{out}, y_{out}]$ 为最终输出的目标点坐标; T_d 为人
为设定的误差阈值, 文中取 $T_d = 64$ 。若 LSTM 估计的
目标点坐标存在负值, 或与预处理结果中的任意可疑
点目标的距离误差超过阈值 T_d , 则认为该时刻图像中
不存在目标; 否则, 取预处理结果中与 LSTM 估计目
标点距离最近的可疑目标点作为最终计算的目标点坐
标。

在 LSTM 估计阶段, 根据可疑目标位置信息序
列, 学习并提取目标的运动特征, 检测目标运动轨迹;
以 LSTM 估计结果为依据, 在预处理结果中选取可信
的目标点坐标, 提升定位精度。

3 实验结果与分析

3.1 仿真算例设计与评价指标

为验证算法对天基预警任务中高动态红外弱小
目标的检测能力, 设计仿真算例序列图像。考虑天基
预警任务的应用场景, 根据低轨探测卫星相机性能及
各背景物体红外特性, 计算生成包含地面、海洋及低
速运动云层的地物背景仿真图像; 在此基础上, 根据
目标高动态特征, 考虑目标可能出现的非线性运动场
景, 向序列背景图像中叠加匀速直线运动、正弦机动
及复杂机动轨迹的目标信号; 为验证算法处理目标出
现和消失场景的能力, 在数据集中设计了相应的仿真
序列图像; 最后, 为提高仿真算例图像的拟真度, 向其
添加标准差为 32 的高斯噪声, 得到平均信杂比为 4.96、
平均信噪比为 4.07 的序列图像。如图 5 所示。训练
集规模为 95 040 张, 测试集 4 950 张。

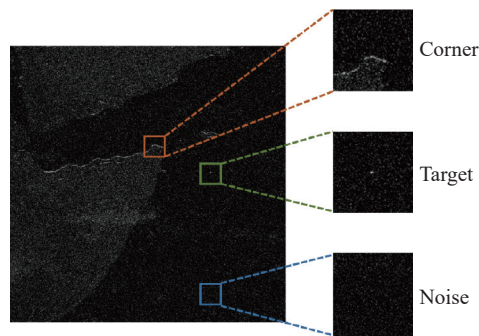


图 5 仿真图片

Fig.5 Simulation image

图 5 为包含目标、背景及噪声的算例图片, 序列
图像中缓慢运动的背景角点及无规律出现的噪声可
能对目标检测造成干扰。对于单目标检测任务中
的一帧图像, 将算法给出的检测结果与真实目标进
行比较, 参考文献 [17] 中的定义方法, 分别使用 TP、FP、
FN、TN 代表可能出现的四种情况:

- 1) 该帧中目标存在且算法认为目标存在, 并计
算出其正确坐标, 即检测结果为正样本, 实际也为正
样本, 定义该情况为 TP (True Positive);
- 2) 该帧中算法认为目标存在, 但所计算的坐标
处无真实目标, 即检测结果为正样本, 实际为负样
本, 定义该情况为 FP (False Positive);
- 3) 该帧中目标存在, 但算法认为目标不存在而
未能计算出其正确坐标, 即检测结果为负样本, 实
际为正样本, 定义该情况为 FN (False Negative);
- 4) 该帧中目标不存在, 且算法正确地认为其不
存在, 即检测结果为负样本, 实际为负样本, 定义
该情况为 TN (True Negative)。

各情况如表 1 所示。

文中使用精确率 (Precision) 和召回率 (Recall) 作
为评估算法的指标^[11]。其计算方法如公式 (13) 所示:

表 1 目标检测中可能出现的情况

Tab.1 Possible situations in target detection

	Detection result: Positive	Detection result: Negative
Reality: Positive	TP	FN
Reality: Negative	FP	TN

$$Precision = \frac{n_{TP}}{n_{TP} + n_{FP}} \in [0, 1]$$

$$Recall = \frac{n_{TP}}{n_{TP} + n_{FN}} \in [0, 1]$$
(13)

式中： n_{TP} 、 n_{FP} 、 n_{FN} 分别表示情况中 TP、FP、FN 出现的数量。即精确率代表算法所做出的检测定位结果中正确的概率，召回率代表目标存在时算法能够正确检测定位的概率，两者值越高，算法的性能越好。

针对不同运动规律的目标，设计并测试了两类算例：实验一中目标坐标以正弦规律变化，用于验证算

法处理较简单的非线性运动目标的能力；实验二中以多正弦波叠加后的函数描述目标坐标变化规律，用于验证算法处理复杂运动规律目标的能力。所有程序运行在相同的硬件平台上，其 CPU 为 i7-11700，GPU 为 Nvidia GTX1660；运行深度学习程序所涉及的显卡驱动程序为 Nvidia-smi510.54，CUDA 版本 11.6。

3.2 实验一

该实验中目标进行简单的非线性运动，运动规律为：

$$x = 256 + 128 \sin \frac{2\pi t}{407}$$

$$y = 256 + 128 \sin \frac{2\pi(t - 31)}{253}$$

式中： x 、 y 均为像素坐标值； t 为帧序号。算法运算结果如图 6 所示。

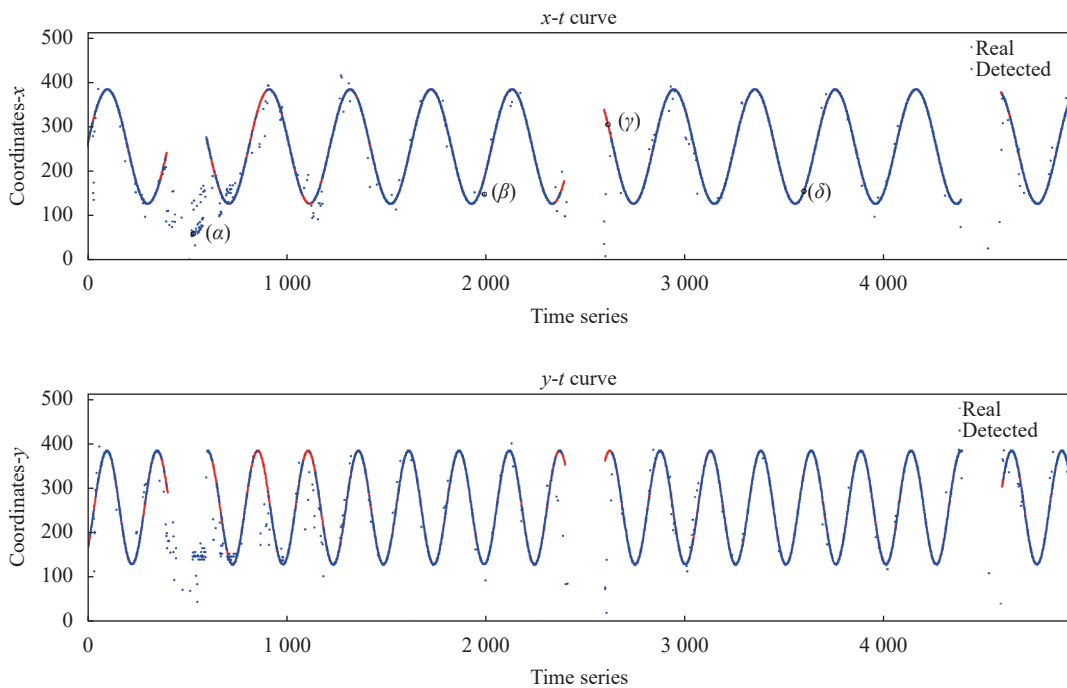


图 6 LISTD 对正弦运动目标的检测结果

Fig.6 LISTD detection results of sinusoidal moving target

在图 6 中，LISTD 算法目标检测精确率为 0.9347，召回率为 0.8851，即在大部分时刻对目标存在性与位置坐标做出了正确估计；目标出现后的 0~50 帧内，算法检测精度低于其他时刻，是由于算法的有效检测需要一定长度的图像序列输入。图 6 中，(α)、(β)、(γ)和 (δ)处的检测场景分别如图 7(a)、(b)、(c)、(d)所示。

图 6 表明大部分时刻算法能够进行如图 7(d)所示的准确检测，但在目标突然出现、噪声及背景显著的情况下，算法有可能出现如图 7(a)~(c)所示的虚警、错检、漏警等错误检测情况。图 7(a)中目标未出现，算法将运动的云层背景角点识别为目标；图 7(b)中算法预测到目标的大致区域，但将附近的噪声信号识别

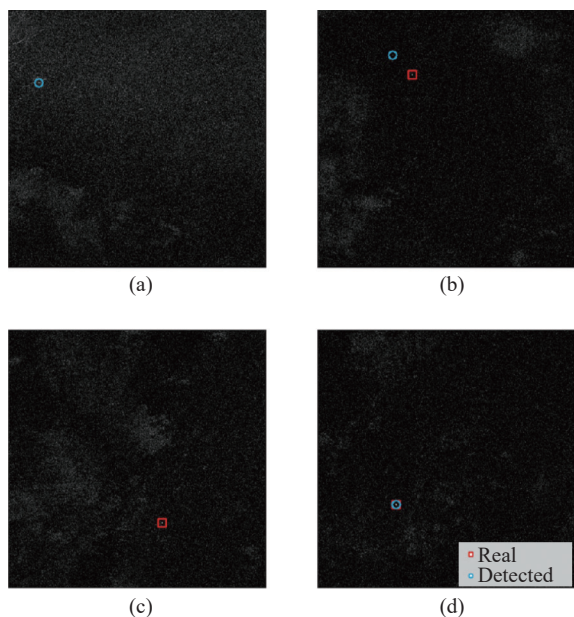


图 7 检测结果。(a) 虚警场景; (b) 错检场景; (c) 漏警场景; (d) 正确检测场景

Fig.7 Detection result. (a) False alarm scene; (b) Mistaken detection scene; (c) Missing alarm scene; (d) Correct detection scene

为目标; 图 7(c)中算法未能在目标出现后的初始阶段进行准确检测; 图 7(d)中算法成功检测出目标并获得准确的位置坐标。

该实验说明运动背景角点、强噪声及短暂出现的

目标有可能限制算法的检测能力,但也证明了在多数情况下算法能够有效完成较强噪声环境下的较简单非线性运动目标的检测任务。

3.3 实验二

为进一步证明文中所提出的 LISTD 算法的有效性,实验二分别在直线、实验一中的典型正弦运动曲线与复杂机动曲线形成的数据集上,与中值滤波算法 (Median Filtering, MF)、MMF^[1]、SHT^[7] 及 YOLOv5^[18] 进行对比,量化分析比较各算法的检测效果。实验中直线运动指目标在图像坐标系中进行速率处于 1.4~5.6 pixel/帧之间的匀速直线运动,在序列图像中,若一条轨迹上的目标超出图像区域,则以前述速率范围内的随机速率在随机方向生成一条新的匀速直线运动轨迹;复杂机动曲线的描述函数由三个周期、幅值不同的正弦型函数叠加得到,如公式 (14) 所示:

$$\begin{aligned}
 x &= 256 + 128 \sin \frac{2\pi t}{407} + 64 \sin \frac{2\pi(t-55)}{193} + 32 \sin \frac{2\pi(t-12)}{202} \\
 y &= 256 + 128 \sin \frac{2\pi(t-31)}{253} + 64 \sin \frac{2\pi(t-17)}{351} + \\
 &\quad 32 \sin \frac{2\pi(t-77)}{189}
 \end{aligned}
 \tag{14}$$

实验结果如表 2 所示。

表 2 算法性能对比

Tab.2 Performance comparison among algorithms

	Uniform linear motion		Sinusoidal motion		Complex motion		Average time consumption/帧·s ⁻¹	Peak memory usage/MiB	Neural network GFLOPs
	Precision	Recall	Precision	Recall	Precision	Recall			
MF	0.6444	0.8311	0.6387	0.8343	0.6415	0.8374	0.007253	360.44	-
MMF	0.4513	0.8715	0.4493	0.8677	0.4510	0.8703	2.7241	359.24	-
SHT	0.9954	0.7404	0.9630	0.7093	0.9585	0.6804	0.009351	359.25	-
YOLOv5	0.3660	0.0949	0.3720	0.1010	0.3400	0.0924	0.06760	597.39	16.422
LISTD	0.9483	0.9237	0.9347	0.8851	0.9362	0.8633	0.008326	1214.13	0.108

从表 2 中可以看出,在精确率方面,对于匀速直线、正弦和复杂机动轨迹的红外弱小目标检测,中值滤波、最大中值滤波、YOLOv5 均低于文中所提出的 LISTD 算法,序列假设检验算法精确率略优于 LISTD,二者没有显著差异,但序列假设检验召回率明显低于 LISTD。在召回率方面,LISTD 高于中值滤波、最大中值滤波、序列假设检验和 YOLOv5 算法。在帧均

耗时方面,LISTD 算法仅次于中值滤波算法,无明显差异,但在精确率和召回率方面,LISTD 均高于中值滤波算法。最大中值滤波由于循环计算过多,导致在 Python 环境中帧均耗时达 2.7 s 左右;YOLOv5 无法提取到目标的纹理、形状特征,在较强噪声的干扰下,漏警及虚警出现率极高,且其网络的复杂性造成了较其他算法更高的时间消耗。在计算资源使用方面,

LISTD 虽然内存占用峰值较高,但该量级不会对运算平台造成较大压力,且由于保持了 LSTM 的轻量级结构,LISTD 的神经网络模型 FLOPs 值远低于 YOLOv5。整体而言,虽然文中所提出的 LISTD 算法在精确率、帧均耗时和内存占用方面不是最优的,但考虑精确率、召回率、帧均耗时和计算资源使用等四方面的综合性能,LISTD 优于对比的其他四种算法。

为测试 LISTD 算法的适用范围,分别在信噪比与信杂比不同的直线、实验一中的典型正弦运动曲线与复杂机动曲线形成的数据集上进行实验,通过在序列仿真图像生成时调整背景图像系数与高斯噪声的标准差,可得到不同的 SCR 与 SNR 组合,实验结果如表 3 所示。

表 3 LISTD 算法对不同信杂比和信噪比数据集的检测结果

Tab.3 LISTD detection results for datasets with different SCR and SNR

SCR	SNR	Uniform linear motion		Sinusoidal motion		Complex motion	
		Precision	Recall	Precision	Recall	Precision	Recall
2	2	0.6905	0.3418	0.7322	0.3549	0.6542	0.3777
	4	0.8191	0.6014	0.7974	0.6168	0.8433	0.6145
	6	0.8674	0.6269	0.8286	0.6430	0.8415	0.6547
4	2	0.8460	0.6131	0.8842	0.6699	0.8709	0.6200
	4	0.8723	0.7498	0.8958	0.7799	0.8740	0.7264
	6	0.9095	0.7845	0.9399	0.8082	0.8507	0.7657
6	2	0.8971	0.7448	0.9229	0.7563	0.8778	0.6653
	4	0.9290	0.8752	0.9500	0.8865	0.9282	0.8577
	6	0.9487	0.8851	0.9454	0.9263	0.9557	0.9161

从表 2 中可以看出,LISTD 适用于较低信噪比与信杂比下高动态非线性运动弱小目标的检测。若序列图像平均信杂比、平均信噪比均高于 2,对于所测试的三种运动规律的目标,LISTD 检测精确率高于 0.6542,召回率高于 0.3418,随着数据集信杂比与信噪比的提高,LISTD 的检测精确率与召回率也相应增加。且经过实验验证,图像平均信杂比为 4.96、平均信噪比为 4.07 时,若目标在长度 128 帧的图像序列中平均速率在 1~6 pixel/帧之间,LISTD 的精确率高于 0.8026,召回率高于 0.5051。实验证明了算法可适用于上述范围内的检测场景。

4 结束语

文中针对天基红外预警任务场景中高动态红外弱小目标检测问题,提出了一种可用于非线性运动目标检测的 LISTD 算法。算法中首先设计了基于中值滤波的自适应预处理模块,将图像信号转化为矢量信息,从而使 LSTM 能够适用于图像问题,扩展了 LSTM 应用场景;然后,利用 LSTM 学习并提取高动态目标

的运动特征,根据可疑目标点位置信息序列检测目标轨迹;最后,设计后处理算法,以 LSTM 估计结果为依据从预处理结果中选取可信的目标点,提升目标定位精度。在自研的红外序列图像数据集上,通过与其他算法的对比,证明了文中算法的有效性和优越性。

文中提出的 LISTD 算法虽然在精确率、召回率和时间消耗方面表现出较好性能,但仍需进一步完善。未来,作者将致力于降低算法检测所需图像序列的长度,开展多目标检测研究,并搭建半实物仿真系统进一步验证算法有效性。

参考文献:

- [1] Deshpande S D, Er M H, Venkateswarlu R, et al. Max-mean and max-median filters for detection of small targets[C]//Proceedings of SPIE, 1999, 3809: 74-83.
- [2] Chen C L P, Li H, Wei Y T, et al. A local contrast method for small infrared target detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(1): 574-581.
- [3] Chen Y, Zhang G, Ma Y, et al. Small infrared target detection based on fast adaptive masking and scaling with iterative

- segmentation [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 9: 7000605.
- [4] Lu X F, Bai X F, Liu J M, et al. Infrared small target detection method based on the improved weighted enhanced local contrast measurement [J]. *Infrared and Laser Engineering*, 2022, 51(8): 20210914. (in Chinese)
- [5] Falconer D G. Target tracking with the hough transform[C]// 1977 11th Asilomar Conference on Circuits, Systems and Computers, 1977. Conference Record. IEEE, 1977: 249-252.
- [6] Barniv Y. Dynamic programming solution for detecting dim moving targets [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1985, 21(1): 144-156.
- [7] Blostein S D, Huang T S. Detecting small, moving objects in image sequences using sequential hypothesis testing [J]. *IEEE Transactions on Signal Processing*, 1991, 39(7): 1611-1629.
- [8] Zhang P, Zhang L, Wang X, et al. Edge and corner awareness-based spatial-temporal tensor model for infrared small-target detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 59(12): 10708-10724.
- [9] Pang D, Shan T, Ma P, et al. A novel spatiotemporal saliency method for low-altitude slow small infrared target detection [J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 19: 7000705.
- [10] Uzair M, Brinkworth R, Finn A. A bio-inspired spatiotemporal contrast operator for small and low-heat-signature target detection in infrared imagery [J]. *Neural Computing & Applications*, 2021, 33(13): 7311-7324.
- [11] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2023-01-05]. <https://arxiv.org/abs/2004.10934>.
- [12] Hou Q, Wang Z, Tan F, et al. RISTDnet: Robust infrared small target detection network [J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 19: 7000805.
- [13] Dai Y, Wu Y, Zhou F, et al. Attentional local contrast networks for infrared small target detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(11): 9813-9824.
- [14] Jiang X, Cai W, Yang Z, et al. Infrared dim and small target detection based on YOLO-IDSTD algorithm [J]. *Infrared and Laser Engineering*, 2022, 51(3): 20210106. (in Chinese)
- [15] Baird Z, Mcdonald M K, Rajan S, et al. A CNN-LSTM network for augmenting target detection in real maritime wide area surveillance radar data [J]. *IEEE Access*, 2020, 8: 179281-179294.
- [16] Liu X, Li X, Li L, et al. Dim and small target detection in multi-frame sequence using Bi-Conv-LSTM and 3D-Conv structure [J]. *IEEE Access*, 2021, 9: 135845-135855.
- [17] Powers D M W. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation[EB/OL]. (2020-10-11)[2023-01-05]. <https://arxiv.org/abs/2010.16061>.
- [18] Jocher G. Comprehensive guide to ultralytics YOLOv5 [EB/OL]. [2023-01-05]. <https://docs.ultralytics.com/yolov5/>.

Research on LSTM method of high dynamic dim and small targets detection for space-based infrared early warning

Zhai Guang, Hu Shengran, Sun Yiyong*

(School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China)

Abstract:

Objective Infrared target detection technology has the advantages of large observation range, short scanning period and strong anti-interference ability, which is of great significance in space-based early warning missions. The infrared targets in space-based early warning missions have the characteristics of small size and weak signal, and it is difficult to use their texture and shape features for detection. In scenes with strong noise and clutter interference, it is necessary to use target motion features for detection. The traditional target detection algorithm generally assumes that targets move with constant velocity, which is not capable of detecting high dynamic nonlinear moving targets with unknown motion rules, and effective detection algorithms for nonlinear moving targets still need further development. To solve the problem of high dynamic and nonlinear moving target

detection in space-based early warning missions, an infrared dim and small target detection algorithm based on LSTM is proposed.

Methods Firstly, an adaptive preprocessing block is proposed, which can extract the location information of suspicious targets and solves the problem of mismatch between LSTM network structure and sequence images.

At the same time, some spatial characteristics of target signals are discarded in order to reduce the amount of computation. Then, due to the problem that traditional algorithms are unable to detect nonlinear motion trajectory, a target trajectory detection method based on LSTM is designed to realize the detection of high dynamic nonlinear moving targets in sequence images. Finally, a post-processing algorithm, which selects target points from the pre-processing results based on the LSTM estimation results is designed to improve the target positioning accuracy.

Results and Discussions Aiming at the targets with different motion laws, two experiments are designed and carried out based on the image sequence with an average signal-to-noise ratio of 4.07. In the first experiment, the target performs simple nonlinear motion with a sinusoidal law. This experiment proves that the proposed algorithm can correctly estimate the existence and position coordinates of the target with a precision rate of 0.934 7 and a recall rate of 0.885 1 (Fig.6-7). In the second experiment, the target performs complex nonlinear motion described by the superposition of multiple sine functions, the proposed algorithm achieves a precision rate of 0.936 2 and a recall rate of 0.863 3 at a detection speed of 0.008 326 s/frame and a peak memory usage of 1 214.13 MiB (Tab.2). Compared with four algorithms including Sequential Hypothesis Testing, the proposed algorithm is proved to have better performance in terms of precision rate, recall rate and detection speed. Through some additional experiments, it is proved that the algorithm is applicable to the detection task in the scenes where the average speed of target is between 1-6 pixels/frame and the average SCR and SNR of sequence images are both higher than 2.

Conclusions Experiments results demonstrate that the proposed algorithm has the ability to detect dim and small targets with high dynamic nonlinear motion from noise and clutter. The proposed algorithm makes use of the spatial and temporal features of the target in the sequence images, and is able to detect targets in low signal-to-noise ratio scenes. And at the same time, with the help of LSTM's ability to extract temporal features, the proposed algorithm is capable in the task of nonlinear moving targets detection. In addition, as the lightweight structure of the LSTM is maintained, the proposed algorithm shows high real-time performance.

Key words: space-based early warning; target detection; LSTM; high dynamic