

低空复杂场景红外弱小目标快速精准检测*

周海^{1,2}, 李保权^{1,2}, 王怀超³, 卞春江^{1,2}

(1. 中国科学院国家空间科学中心 复杂航天系统电子信息技术重点实验室, 北京 100190;

2. 中国科学院大学, 北京 100049; 3. 中国民航大学 计算机学院, 天津 300300)

摘要:针对低空复杂场景下红外弱小目标检测难度大、虚警率高等问题,面向探测系统中高帧频图像实时处理应用需求,提出基于全卷积网络的弱小目标精准检测方法和基于现场可编程逻辑门阵列(field programmable gate array, FPGA)的低时延并行处理方法。采用轻量化全卷积网络对红外图像中弱小目标进行空域检测,对相邻图像帧疑似目标进行时域轨迹关联以进一步降低虚警率。实验结果表明:上述方法相比于五种传统方法在检测率和虚警率性能方面均有显著提升,并在单片FPGA上完成100 Hz图像实时处理,处理时延低于1.8 ms,实现低空复杂场景弱小目标高精度高鲁棒快速实时检测。

关键词:红外弱小目标;目标检测;全卷积网络;实时处理;现场可编程逻辑门阵列

中图分类号:TN216 **文献标志码:**A **文章编号:**1001-2486(2023)01-074-12

Fast and accurate detection of infrared dim small target in low altitude complex scenes

ZHOU Hai^{1,2}, LI Baoquan^{1,2}, WANG Huaichao³, BIAN Chunjiang^{1,2}

(1. Key Laboratory of Electronics and Information Technology for Complex Space Systems, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Department of Computer, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Aiming at the high difficulty and high false alarm rate of infrared dim small moving target detection in low altitude and complex scenes, a precise detection method based on full convolution network and a low-latency parallel processing method based on FPGA (field programmable gate array) were proposed to meet the real-time processing application requirements of high frame rate images in the detection system. A lightweight full convolutional network was used to detect dim small targets in spatial domain, and time domain trajectory correlation on suspected targets in adjacent frames was performed to further reduce false alarms. The experimental results show that the detection rate and false alarm rate are significantly improved compared with the five traditional methods. And the real-time processing of 100 Hz image is completed on a single FPGA with the processing delay less than 1.8 ms. This method realizes high-precision, high-robust and fast real-time detection of dim small targets in complex low altitude scene.

Keywords: infrared dim small target; target detection; full convolution network; real-time processing; field programmable gate array

基于红外传感器的低空运动小目标探测因具有隐蔽性高、覆盖范围广和全天候工作的特点,在现代国防、红外监视和空中交通管制等军民领域均发挥着重要作用^[1-3]。真实低空场景下,红外小目标检测依然受到诸多挑战:①由于观测距离远,无人机、飞机模型、风筝等低空目标往往很小很弱,形状和纹理信息不明显;②由于低空观测环境复杂,背景中可能含有树木、植被、地物等强干扰^[4](如图1所示),这些干扰源进一步加大了小目标检测难度;③红外传感器在探测跟踪系统中

主要定位于目标发现,在处理中需以极低的时延生成检测信息用于对系统内其他单元进行引导,检测性能与处理实时性是决定系统整体性能的关键性设计。

近20年来,随着红外探测器的快速发展与应用,出现了大量针对各种场景的红外小目标检测方法。这些方法主要分为两类:基于单帧的检测方法和基于序列的检测方法。基于单帧的检测方法通常利用图像空间域的局部差异^[5]设计不同的滤波器来抑制噪声和杂波,如top-hat变换^[6]、

* 收稿日期:2021-03-24

基金项目:国家部委基金资助项目(1916318ZD00400201);中国科学院青年创新促进会资助项目(E0293401)

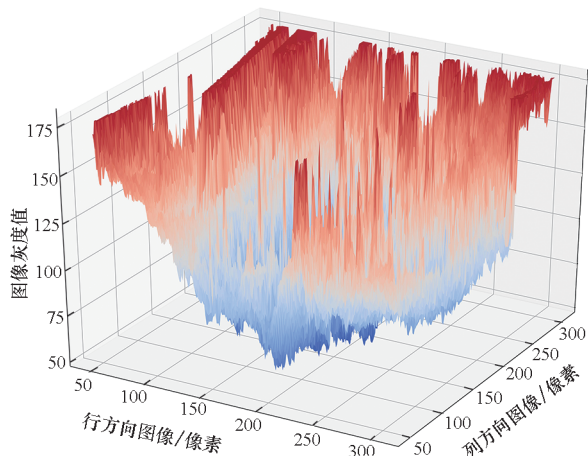
作者简介:周海(1987—),男,安徽蚌埠人,副研究员,博士研究生,E-mail:zhouhai@nssc.ac.cn;

李保权(通信作者),男,河南罗山县人,研究员,博士,博士生导师,E-mail:lbq@nssc.ac.cn



(a) 低空复杂场景代表性目标

(a) Low-altitude complex scenes representative targets



(b) 低空复杂场景能量三维图示

(b) Low-altitude complex scenes energy 3D diagram

图1 低空复杂场景代表性目标及相应的能量三维图示

Fig. 1 Representative targets in low-altitude complex scenes and corresponding energy 3D diagram

最大中值和最大均值滤波器^[7]等。文献[8]基于人类视觉注意机制提出了局部对比度量(local contrast measure, LCM)方法。在此方法基础上,产生了许多改进的LCM方法^[9],如多尺度相对局部对比度量(relative local contrast measure, RLCM)^[10]和基于多尺度面片的对比度量(multiscale patch-based contrast measure, MPCM)^[11]等。对于简单场景,这些方法可以有效抑制杂波并增强目标,但对于低空复杂场景,由于存在大量与目标相似的干扰导致检测虚警率较高。基于序列的检测方法利用时域相关性对图像背景进行估计,并利用空域信息对目标进行检测。时空局部对比度滤波器(spatial-temporal local contrast filter, STLCF)使用空间局部对比度和时间局部对比度来增强目标^[12]。此外,许多方法使用稀疏表示来

分离连续图像上的背景和目标^[4]。时间滤波和关联策略(temporal filtering and association strategy, TFAS)方法利用最大滤波和中值滤波的时间融合(max filter and median filter temporal fusion, MMTF)来检测目标,然后执行关联操作以形成连续的轨迹^[13]。与基于单帧的检测方法相比,基于序列的检测方法使用了额外的图像时间维度信息,其检测精度更高。

近些年,深度学习方法飞速发展,在通用目标检测领域产生了革命性的进步,针对具有丰富纹理特征的大尺度目标,产生了很多深度检测网络,如FasterRCNN^[14]、YOLO^[15]、SSD^[16]等。但是,针对低空红外弱小目标,由于缺乏显著的纹理和形状信息,这些网络检测性能并不理想。

针对低空复杂场景下红外弱小目标检测率和虚警率方面的不足,本文在传统方法基础上引入全卷积网络(fully convolutional network, FCN)进行性能提升。

1 检测算法原理

本文算法充分利用时空域信息实现弱小目标高精度检测,整体流程如图2所示。对于每一帧图像首先采用全卷积网络得到目标检测概率分布热图,并在阈值分割后通过空域连通将杂散疑似目标连通标记成备选关联目标。由于目标检测概率分布热图中存在的虚警一般由噪声、地上固定物或盲元产生,通常不具备规则的运动信息,故可以将相邻两帧备选关联目标进行轨迹关联,进一步消除虚警,得到鲁棒的检测结果。

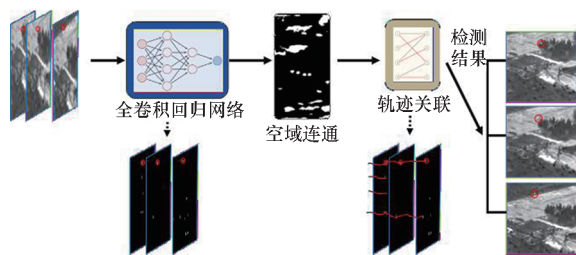


图2 本文算法流程

Fig. 2 Framework diagram of the proposed method

1.1 全卷积网络

红外弱小目标由于没有显著的轮廓和形状信息,很难直接应用当前的深度检测网络。低空弱小目标在远距离情况下会退化成尺度非常小的目标,因而需要预测一幅图像中所有像素点的类别,是个像素密集型检测任务。全卷积网络^[17]能够对图像进行像素级的分割,其网络结构设计思想非常适用于红外弱小目标检测任务。同时该网络

具有以下几方面优势:①输入图像可以是任意尺寸,因而图像处理时不需要特定步骤对其进行尺寸预处理,并且当需要对图像分块以满足并行流水化处理要求时,可对图像进行任意大小的块分割,便于匹配并行处理架构计算需求;②对目标的分割分类属于像素级的分类方法,并且具备尺度和旋转不变性;③每个像素对应位置像素点的端对端训练方式比利用图像上下文信息和局部特征训练参数更加高效。

但传统全卷积网络是针对通用实例分割任务设计,通常分割的目标较大,所以不加改进直接应用于红外弱小目标检测会影响检测性能。为提高网络的红外弱小目标检测精度,权衡硬件实现成本,本小节对全卷积网络结构进行了优化,构建了一种适合红外弱小目标空域检测的全卷积回归网络。

1)减少网络层数:卷积神经网络原理上是通过更深的网络层数来学习更抽象的特征表示,这样可以有效增加通用场景的目标检测精度。但对于低空红外图像,由于图像噪声强度大,小目标形状和纹理特征非常弱,所以较深的网络层数很难学习到有效的抽象特征表示,进而影响检测精度。对传统全卷积网络的 16 个卷积层进行削减,一方面保持对小目标底层特征的充分利用,另一方面随层数降低可以有效减小模型参数,降低计算量。经过多次实验,当网络层数逐步减少至 5 时,网络精度变化小于 0.3%;当网络层数减少至 4 时,精度下降明显。权衡精度与计算量,算法采用 5 层卷积层构建检测网络。

2)降低池化层影响:池化层步长越大,则该层之后的其他各层存储计算资源消耗越小,但过大的步长也会损失图像细节特征。为了确保网络检测精度并得到稠密的检测结果,经多次实验,仅在将第一个池化层的步长设置为 2、其余池化层的步长均设置为 1 的情况下,可以在不影响检测精度的前提下降低部分硬件资源消耗。

3)使用回归头:传统全卷积网络输出层经过 softmax 函数得到分类结果,而对于小目标检测,如果将其最终结果简单看成是二分类问题,那么正负样本会严重失衡,且不能得到小目标的精确位置。因此输出层使用均方误差回归损失函数替代 softmax 函数,可以得到当前像素点 x 是目标的概率 $p(t|N_x, x_j \in N_x)$,其中 N_x 表示像素点 x 的邻域。

4)优化归一化(batch normalization, BN)层位

置:归一化层主要用于网络训练过程中加快训练收敛,避免梯度消失和过拟合。传统全卷积网络中归一化层通常放在激活层后面,但归一化层仅作用于网络训练阶段,对于网络推理阶段其参数已经固定,适合与卷积层通过融合处理的方式降低计算量。将其调整至卷积层与激活层中间,加速网络训练的同时又充分考虑实时处理中的优化需求。

通过采用以上改进,全卷积回归网络结构参数如表 1 所示。

表 1 网络参数
Tab. 1 Network parameter

层编号	输出 通道数目	输入 通道数目	卷积 核尺寸	步长
Conv1	16	1	3 × 3	1
BN1	16	16		1
ReLU1	16	16		1
MaxPool	16	16		2
Conv2	32	16	3 × 3	1
BN2	32	32		1
ReLU2	32	32		1
Conv3	32	32	3 × 3	1
BN3	32	32		1
ReLU3	32	32		1
Conv4	32	32	4 × 4	1
BN4	32	32		1
ReLU4	32	32		1
Conv5	1	32	1 × 1	1

1.2 轨迹关联

轨迹关联利用目标多维特性进行时序关联性度量,提取高疑似度目标关联序列,抑制虚假目标关联序列的产生,从而实现对目标低虚警率检测。轨迹关联前需将相邻的疑似点连通成一个目标区域,计算目标特性,包括位置、区域大小、能量等,利用这些目标特性度量历史轨迹与当前目标的关联度,实现时域上的轨迹关联。关联过程中未能与历史轨迹相关联的疑似目标均定义为备选起始轨迹供下次关联,各历史轨迹连续多帧不能与新疑似目标关联则判定轨迹消失。关联度量包括:窗口关联度量、能量稳定性关联度量、速度方向一致性关联度量。

1.2.1 窗口关联度量

由于目标运动速度有限,相邻帧间目标的位移也是有限的,因此可以基于历史关联序列位置

设置一定大小的关联窗口,对落入窗口内的疑似目标进行关联。关联窗口的设置方式如图3所示,考虑到目标可能会在某一帧发生漏检,若上一帧未关联数据,则将窗口大小适当外扩以保证对受遮蔽目标的间断关联。图3中 Δx 和 Δy 为序列末点与待关联疑似点的坐标差,若 $\max(\Delta x, \Delta y)$ 小于关联窗口半径,则认为疑似目标满足窗口关联条件。

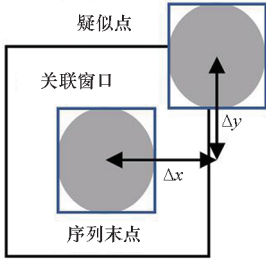


图3 窗口关联示意

Fig. 3 Window association diagram

1.2.2 能量稳定性关联度量

同一目标量的能量在相邻帧之间是相对稳定的,可选取轨迹末端点能量峰值 E_{seq} 和当前疑似点能量峰值 E_{uc} 中的最小值作为参考值,如式(1)所示,计算待关联疑似点能量峰值与序列末点能量峰值绝对变化量,其相对参考值的百分比为能量关联度,若该值小于阈值 T_{eng} ,则认为待关联点满足能量稳定性关联。

$$\frac{|E_{uc} - E_{seq}|}{\min(E_{uc}, E_{seq})} < T_{eng} \quad (1)$$

1.2.3 速度一致性关联度量

速度方向一致性判断如图4所示,由于帧间目标运动速度方向在短时间内变化较小,因此基于速度大小和方向一致性可对疑似目标进行关联筛选。

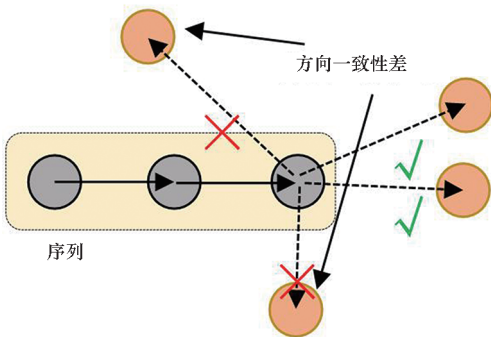


图4 速度方向一致性判断

Fig. 4 Consistency judgment of speed and direction

具体方法通过计算轨迹方向与目标点方向矢量夹角大小来判断是否满足方向一致性条件,通

过计算轨迹速度与目标点速度差值判断是否满足速度一致性条件,并通过阈值进行调整。

2 检测算法优化与实现

本文算法主要面向低空场景小目标红外探测跟踪系统应用,为满足红外探测器高帧频图像处理需求与系统强实时性高稳定性要求,采用现场可编程逻辑门阵列(field programmable gate array, FPGA)进行实现。不同于中央处理器(central processing unit, CPU)、数字信号处理(digital signal processing, DSP)芯片等传统嵌入式处理器,FPGA可以为复杂算法内各处理模块定制单独的硬件处理电路,并针对不同算法模块计算特点优化处理结构,构建低时延并行流水处理架构。本节首先给出算法顶层架构设计分析,继而对算法各模块进行并行化和流水化改进。

2.1 顶层流水架构设计

高帧频红外探测器由于帧周期较短,图像传输时间 P 往往占据帧周期 T 的80%以上,因而为实现高帧频图像快速处理,实时输出每帧处理结果,并降低复杂算法各级模块计算与吞吐压力,需要采用图像分块处理思想构建顶层模块间流水处理架构,如图5所示。

图像按行传输等分为 M 块,在第 N 块图像传输过程中,可同时进行 $N-1$ 块图像全卷积推理处理、 $N-2$ 块图像连通处理和 $N-3$ 块图像轨迹关联处理。当 M 取值大于等于3且各处理模块对块图像处理时间小于等于 T/M 时,可实现各级模块并行流水高利用率实时计算,此时一帧图像处理时延最大值为 $3T/M$,并随着 M 取值增大显著降低,实现快速目标检测处理。

但 M 取值过大时也会对块间数据交互与融合带来额外处理压力,本文对尺寸为512像素×512像素的100 Hz红外图像进行处理时取 $M=16$,此时对于单块图像各级模块处理时间应小于0.625 ms。

2.2 全卷积网络优化与实现

采用的全卷积网络在设计中已经充分减少网络层数和模型参数,从而构建轻量化网络。但相对传统算法,全卷积网络计算规模依然较高,因此需进一步优化网络层次、降低计算复杂度,并充分利用FPGA内逻辑资源、DSP硬核资源与片内缓存资源,构建高效并行计算架构,实现全卷积网络算法快速实时处理。

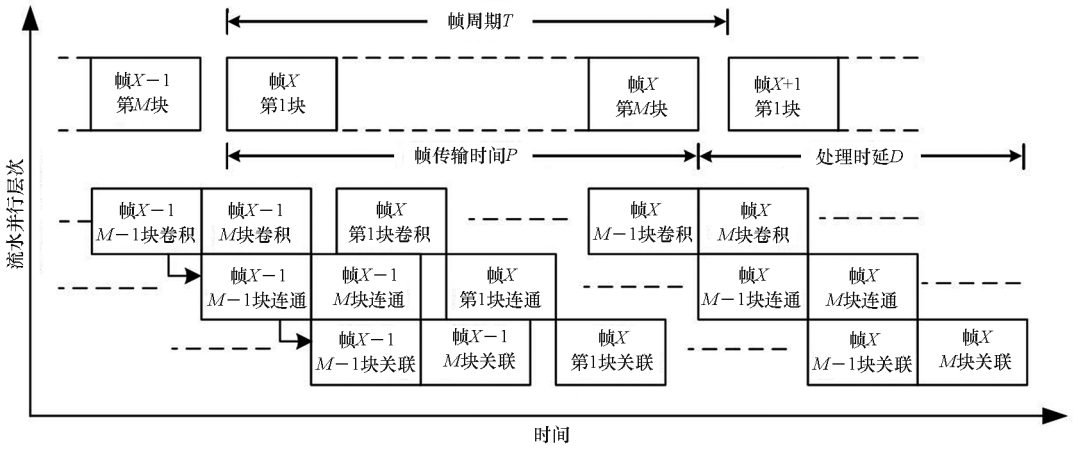


图 5 算法实现顶层流水架构示意

Fig. 5 Schematic diagram of algorithm implementation top-level pipeline architecture

2.2.1 归一化层融合优化

由 1.1 节可知,算法中全卷积网络主要包含卷积层、池化层与归一化层(BN 层)。其中 BN 层主要在网络训练过程中解决梯度消失和梯度爆炸问题,同时提高泛化性能。BN 层的计算公式为:

$$y_{bn} = \gamma \frac{x_{bn} - \mu}{\sqrt{\sigma^2 + \varepsilon}} + \beta \quad (2)$$

式中, x_{bn} 为 BN 层输入, y_{bn} 为 BN 层输出, γ 为尺度因子, β 为偏移因子, μ 和 σ^2 为每个通道的平均值与方差, ε 为方差调整因子。BN 层计算涉及高精度乘加运算,因而在 FPGA 实现时需要消耗大量计算时间与参数存储资源,并会显著降低全卷积网络处理时延。同时由于 BN 层运算与卷积计算模式差别较大,在 FPGA 实现中很难与卷积计算共享一套处理架构,导致 BN 层带来较大计算资源开销。

在网络训练结束后, BN 层各参数在推理过程中均为静态参数,可通过将 BN 层参数融入卷积层参数的方法移除 BN 层计算,显著简化处理流程,具体参数融合方法推导如下。

设卷积计算公式为:

$$Y = W * X + b \quad (3)$$

当卷积层输出为 BN 层输入时,式(3)可在式(2)中展开。

$$y_{bn} = \frac{\gamma}{\sqrt{\sigma^2 + \varepsilon}} W * X + \left[\beta + \frac{(b - \mu)\gamma}{\sqrt{\sigma^2 + \varepsilon}} \right] \quad (4)$$

此时为省略 BN 层计算,直接将 y_{bn} 看成融合后新卷积层输出,则对于新卷积层只需要按照式(5)和式(6)更新卷积层参数即可。

$$W' = \frac{\gamma}{\sqrt{\sigma^2 + \varepsilon}} W \quad (5)$$

$$b' = \beta + \frac{(b - \mu)\gamma}{\sqrt{\sigma^2 + \varepsilon}} \quad (6)$$

由于式(5)和式(6)中所有参数在训练结束后均为静态参数,且精度相同,因此在全卷积网络推理处理前可预先完成参数融合,在保证计算精度的同时直接减除 FPGA 内原 BN 层处理资源消耗。融合后网络结构如表 2 所示。

表 2 BN 层融合后网络结构

Tab. 2 Network structure after BN layer fusion

层编号	输出通道数目 O_c	输入通道数目 I_c	卷积核尺寸 $W_x \times W_y$	输入特征图尺寸 $I_x \times I_y$	输出特征图尺寸 $O_x \times O_y$
Conv1	16	1	3 × 3	32 × 512	32 × 512
ReLU1	16	16		32 × 512	32 × 512
MaxPool	16	16		16 × 256	16 × 256
Conv2	32	16	3 × 3	16 × 256	16 × 256
ReLU2	32	32		16 × 256	16 × 256
Conv3	32	32	3 × 3	16 × 256	16 × 256
ReLU3	32	32		16 × 256	16 × 256
Conv4	32	32	4 × 4	16 × 256	16 × 256
ReLU4	32	32		16 × 256	16 × 256
Conv5	1	32	1 × 1	16 × 256	16 × 256

2.2.2 卷积层并行优化

由表 2 可以看出,经融合优化后的全卷积网络由卷积层、激活层和池化层组成,由于激活层采用 ReLU 函数,其与池化层在处理上仅需采用简单比较,不涉及复杂乘加运算,复杂度与卷积层卷积运算相差较大,因而对全卷积网络计算优化重

点主要聚焦在各卷积层。卷积运算是涉及输入特征图和卷积核权值的三维乘积运算,由 4 个层次的循环实现,如图 6 所示。

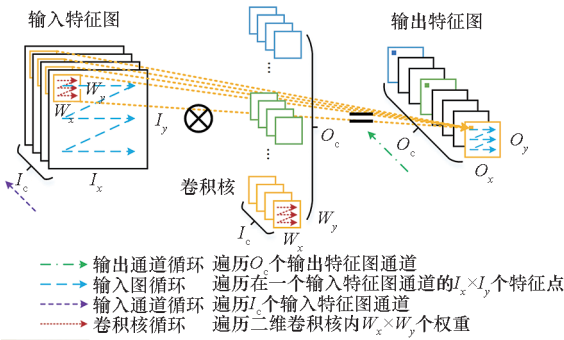


图 6 卷积运算 4 层循环示意图

Fig. 6 Four layer loop diagram of convolution operation

从数学角度来看,在对卷积层进行并行优化设计时,可以任意对一个或多个层次循环进行全部或部分展开。但由于展开的大小受限于 FPGA 片内资源,因而如何选择合适的展开层次与规模是卷积并行优化设计的重点。

算法采用的全卷积网络虽然卷积层数只有 5 层,但卷积核参数依然高达 30 384 个,单层输入特征图点数最大为 131 072, FPGA 片内缓存不足以存储所有数据,需要外部存储器对卷积核参数和层间特征图进行缓存。因此,基于 FPGA 的全卷积网络实时处理单元由 3 个层次的存储结构组成:①外部大容量存储器;②FPGA 片内缓存;③与卷积计算实时处理引擎(processing engine, PE)阵列直接耦合的乘加运算临时寄存器(register, REG)。存储结构如图 7 所示。

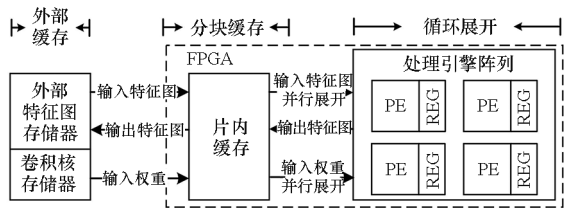


图 7 卷积层实时处理存储结构

Fig. 7 Convolutional layer real-time processing storage structure

每层网络处理时从外部存储器中获取待处理特征图分块和相关卷积核参数到片内缓存,然后按照 PE 阵列规模将数据转发 PE 中的寄存器。在 PE 阵列计算完成后将结果传递回片内缓存,并在必要时存入外部存储器,等待下一层网络计算调用。

为实现全连接网络低延时高效处理,上述处理流程应在 PE 阵列并行计算的基础上构建细分

流水处理架构,即尽可能保证各 PE 在每个时钟周期内都能执行一次运算。循环并行展开方式的不同直接决定了 PE 阵列规模和 PE 数据流模式,继而影响到 3 个层次存储间数据存储访问模式与片内数据复用机会。

此外,在卷积层优化设计时应充分考虑不同卷积层尺寸的差异,提高 PE 阵列并行流水处理架构对不同卷积层的匹配性,以最大限度减少由于各卷积层差异带来的计算效率降低。从表 2 可以看出,网络各卷积层卷积核尺寸差别较大,包含 1×1 、 3×3 、 4×4 三种,因而不适合对卷积核循环进行展开;输入特征图和输出特征图通道数最小均为 1,因此也不适合对输入输出通道进行展开;全卷积网络结构具有良好的输出特征图尺寸一致性特点,有利于输入特征图循环展开。输入特征图循环并行展开如图 8 所示。

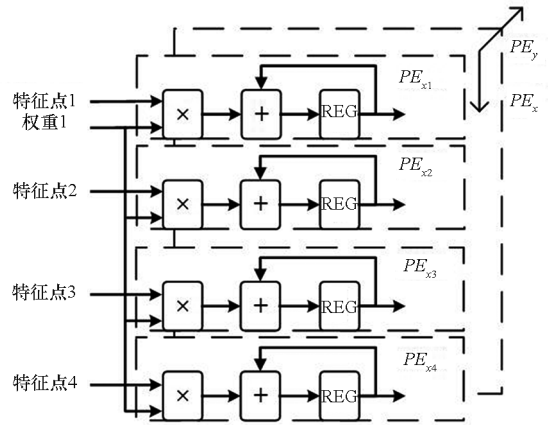


图 8 输入特征图并行展开

Fig. 8 Parallel expansion of input feature map

为最大化 PE 阵列利用率,应将并行展开变量 (PE_x, PE_y) 设置为所有卷积层输入特征图尺寸变量 (I_x, I_y) 的公因子,由表 2 可知网络 (PE_x, PE_y) 最大可取值为 (16, 256),但由于每个 PE 在实现上需消耗 FPGA 内 1 个 DSP 硬核,受限于 DSP 硬核数量,在实际设计中 (PE_x, PE_y) 可取 (16, 128),此时并行展开度和 PE 阵列规模为 2 048。

2.2.3 全卷积网络计算架构设计

基于 BN 层融合与卷积并行展开优化设计,图 9 给出了全卷积网络并行计算架构。架构外围设置了 2 个动态随机存取存储器(dynamic random access memory, DRAM),分别存储特征图和卷积核。每层卷积处理开始时 DRAM 在接口管理模块控制下将特征图分块和相应权重分别加载至片内输入特征图缓存与片内权重缓存中,继而依据 PE 阵列尺寸,通过内部复用数据总线将特征图和权重数据载入各 PE,开始高速并行计算。

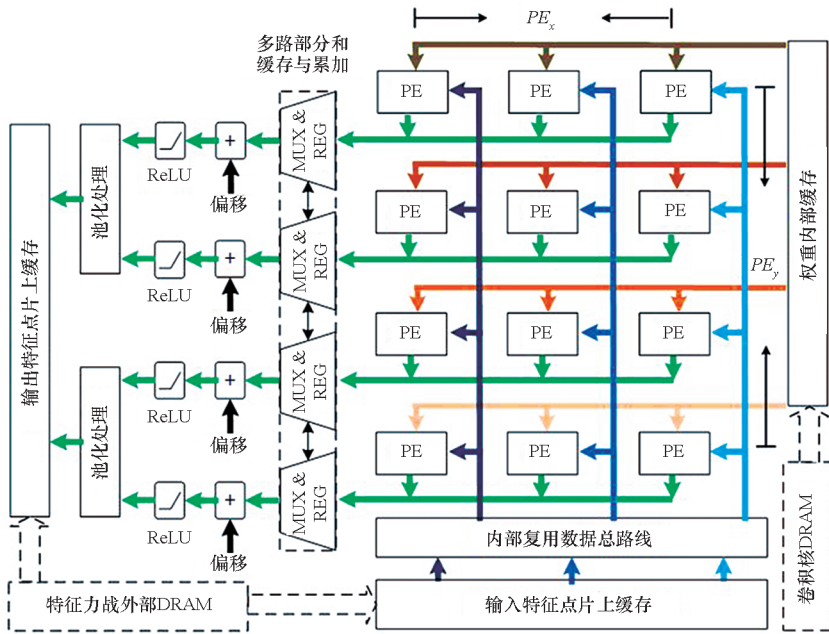


图 9 全卷积网络并行计算架构

Fig. 9 Full convolution network parallel computing architecture

数据在 PE 阵列中按输入特征图展开尺度进行每周周期不间断并行计算。为尽快得到对应局部输出特征图,降低部分和寄存器数量,便于后端 ReLU 与池化流水处理,设计时将输出通道循环放在最后一层,先进行卷积核循环与输入通道循环。同时考虑卷积核尺寸较小,为加强 PE 阵列数据复用降低缓存吞吐压力,设计中先完成卷积核循环输出单通道部分和,继而完成输入通道循环,最后交由多路部分和缓存与累加单元完成数据累加生成。

卷积处理后结果依次传输至偏移累加计算单元、ReLU 计算单元和池化处理单元,最终按序存入外部 DRAM,等待下一层计算读取。这三层处理由于计算规模较小耗费资源较少,在设计中为尽可能降低处理时延,也按特征图列方向进行了不同尺度的并行化设计。

由于每层运算的参数和尺寸是固定的,因而在存储区中采用少量乒乓缓存技术就可以实现 4 个卷积循环迭代并行流水和卷积层间的顺序流水计算。所有缓冲区读写地址和 PE 阵列计算顺序也在编译过程中依据网络模型预先分配并记录,有效降低数据传输额外占用的时间,支撑 PE 阵列高效并行工作,提高全卷积网络并行计算架构处理性能。

2.3 轨迹关联优化与实现

轨迹关联在设计中为实现低时延处理,重点考虑两个层面的优化加速:一是历史轨迹与多个备选点之间的并行处理,二是单次关联过程中不同维度一致性计算间的并行处理。轨迹关联架构

设计如图 10 所示,轨迹关联主要包括序列与疑似点存取、关联计算和序列更新 3 个模块。序列与疑似点存取模块完成对已关联序列和待关联疑似点存取;关联计算模块完成窗口关联、能量一致性关联和速度一致性关联处理;序列更新模块对关联结果融合判断产生新的序列,并输出存入外部缓存等待下次关联。

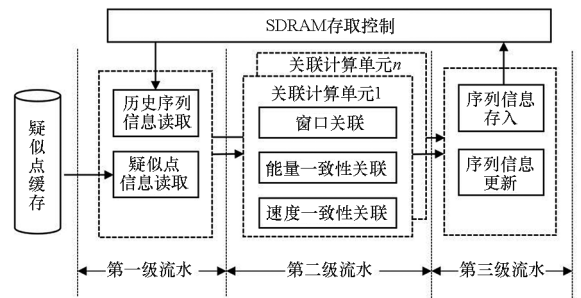


图 10 轨迹关联架构设计

Fig. 10 Architecture design of trajectory association

模块采用流水优化方法提高轨迹关联处理速度。理论上处理性能会随流水的级数增加而提高,但同时也会增大对硬件资源特别是片内有限存储资源的消耗。依据模块划分采用三级流水模式,在第二级流水中通过采用多个关联计算单元支撑多备选点与当前轨迹同步并行关联计算,有效降低多备选点的关联处理时间。此外,不同于常规串行关联判断处理方法,对窗口关联、能量一致性关联和速度一致性关联采用独立并行计算后,融合表决判断,进一步大幅度缩短计算时延、提高处理性能。







3 测试与结果分析

3.1 实验数据集

实验数据集采用由国防科技大学公开的真实低空场景红外小目标数据集^[18],其中包含 22 个红外序列。序列为地面背景或地面与天空混合背景,包含山、树木、建筑、植被、水等干扰物。为了充分验证本文算法对于复杂场景下红外弱小目标的检测性能,选取了其中场景较复杂的 6 个序列用于性能测试和分析,其余序列用于构建训练数据集。训练数据集建立过程中除图像中原有目标,还随机添加了不同位置、信噪比和大小的仿真目标对数据集进行了增广,共计 13 874 幅。表 3 给出了 6 个测试序列的细节描述,共计 6 649 幅红外图像。从目标信噪比范围可以看出场景中的目标为弱目标^[19],各个序列的图像复杂程度用序列中各图像场景信噪比(scene signal noise ratio, SSNR)的平均值进行描述,值越大表明场景越复杂,

表 3 测试序列与图像参数

Tab.3 Test sequence and image parameters

序列号	典型帧	总帧数	平均 SNR	目标信噪比
1		3 000	6.413	[0.97, 8.04]
2		399	6.326	[0.62, 7.52]
3		751	5.030	[1.47, 5.70]
4		1 599	4.168	[0.25, 5.16]
5		400	3.558	[1.25, 4.23]
6		500	6.254	[1.65, 5.83]

图像 SSNR 定义如式(7)所示。

$$SSNR = 10 \lg \left[\frac{\sum_{i=1}^M \sum_{j=1}^N D_{i,j}}{\sum_{i=1}^M \sum_{j=1}^N (S_{i,j} - L_{i,j})} \right] \quad (7)$$

式中, M 和 N 分别代表图像长宽, $D_{i,j}$ 代表图像中坐标为 (i, j) 的点邻域方差, $S_{i,j}$ 代表坐标为 (i, j) 的点中心灰度值, $L_{i,j}$ 代表坐标为 (i, j) 的点邻域均值。

3.2 实验与测试平台

实验与测试平台采用课题组自研高性能实时处理机予以验证,检测算法运行在图像处理板上 FPGA 内,图像处理板实物如图 11 所示。

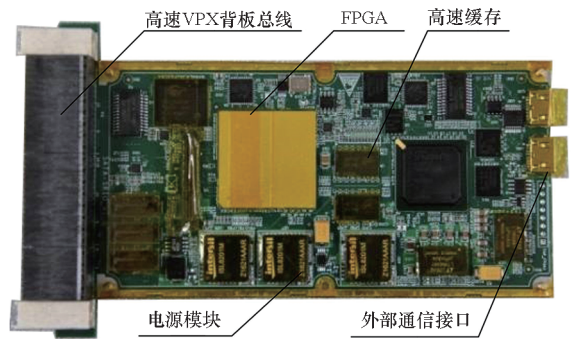


图 11 自研图像处理板

Fig. 11 Self-developed image processing board

为验证算法处理性能,以处理机为中心搭建了一套自测试验证系统,系统组成如图 12 所示。系统采用图像模拟源设备实时高帧频发送测试序列图,由检测结果记录设备实时接收并存储算法各项处理结果,其中包含全连接网络输出的空域检测结果与多帧关联输出的时域检测结果,继而通过结果分析实现对算法功能与性能的测试验证。

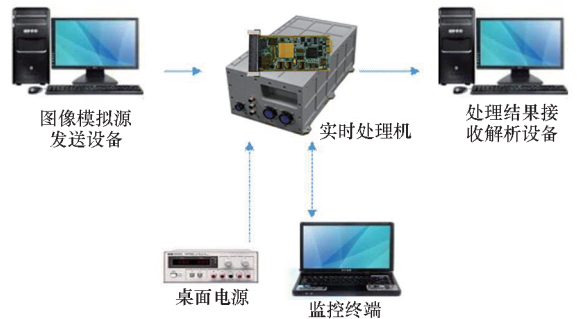


图 12 自测试验证系统

Fig. 12 Self-test verification system

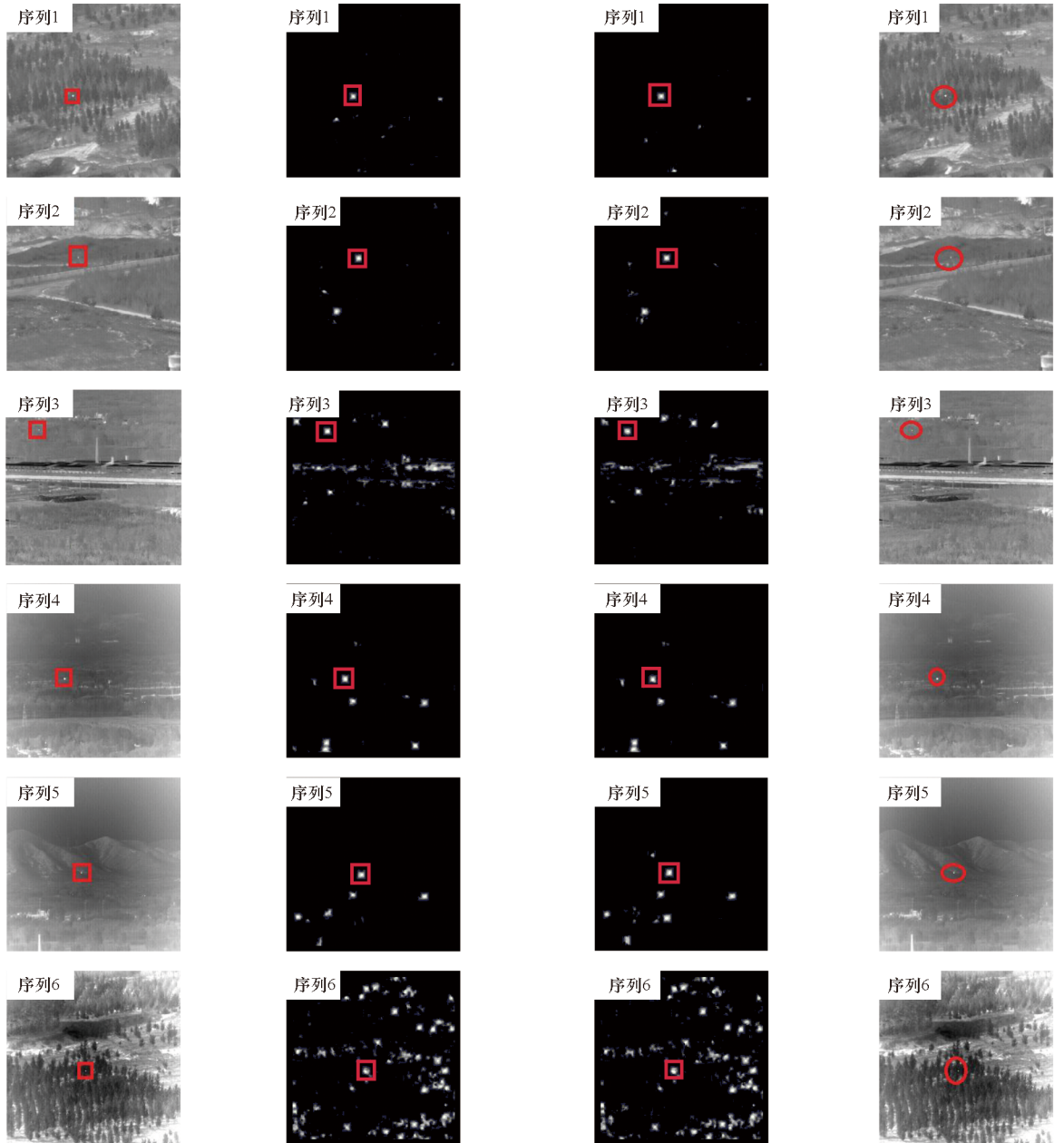
3.3 性能测试与分析

图 13 按行从上至下分别列出了 6 个测试序列中的典型场景,并标注出本文所提算法在

FPGA 处理后获得的检测结果。通过图 13(a) 可以看出,受大量干扰物影响,人眼几乎很难分辨目标位置。图 13(b)和图 13(c)为图 13(a)相邻两帧经过全卷积网络回归得到的目标概率热图,如红框所示位置,目标在热图中均有很强的响应。但由于场景比较复杂,一些噪声和干扰物也有较强响应。图 13(d)为关联后检测结果,可以直观看出,相邻两帧目标概率热图之间存在不同位置的高响应区域,这些响应由随机噪声引起,通过轨迹关联可以有效剔除这一类高响应引起的虚假目

标,抑制虚警,最终得到准确的检测结果。

本文方法与其他 5 个基准方法的接受者操作特性(receiver operating characteristic, ROC)曲线比较如图 14 所示。其中,通过不同的分割阈值,获得了 MPCM、RLCM、STLCF 和 MMTF 4 种基准方法的 ROC 曲线。TFAS 方法的 ROC 曲线通过不同的位移增量阈值获得。本文方法的 ROC 曲线同样通过不同的分割与关联阈值获得。与基准方法相比,本文方法在很低的误报率下,能够获得较高的检测率。表 4 给出了本文方法和其他 5 种



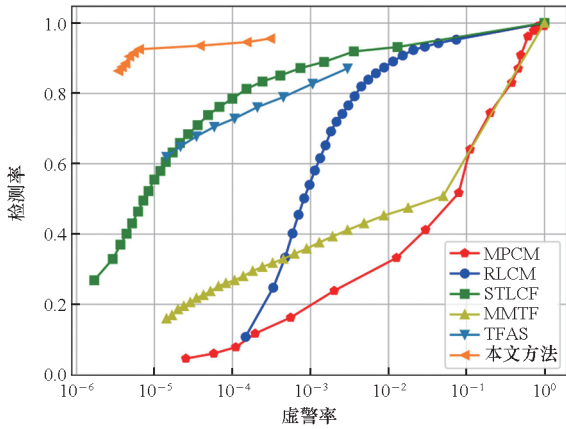
(a) 测试序列典型场景 (a) Typical scenes in test sequence
 (b) 当前帧目标概率热图 (b) Target probability heat map of current frame
 (c) 相邻帧目标概率热图 (c) Target probability heat map of adjacent frame
 (d) 关联后检测结果 (d) Detection result after correlation

图 13 本文算法实现测试结果

Fig. 13 Algorithm implementation test results of this paper

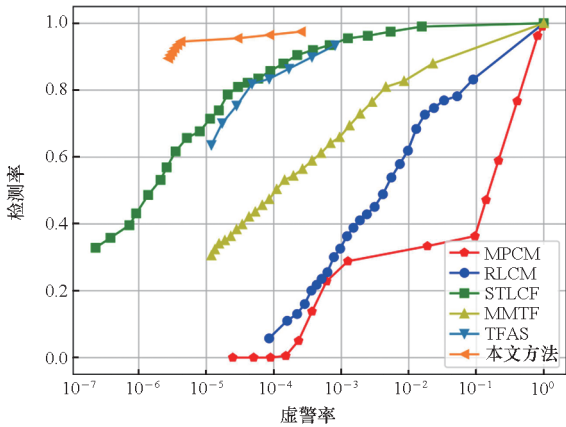
基准方法在 6 个测试序列下的平均虚警率和平均检测率对比。从中可以看出,本文方法在平均虚警率优于基准方法 4 倍以上(本文为 0.000 1)条件下,平均检测率相比基准方法提高 20% 以上(本文为 93%),并对不同场景具有良好的鲁棒性。

同时表 4 给出了测试中运行平均速率,其中 5 个基准方法均在 X86 平台(CPU 为 Intel Core I7



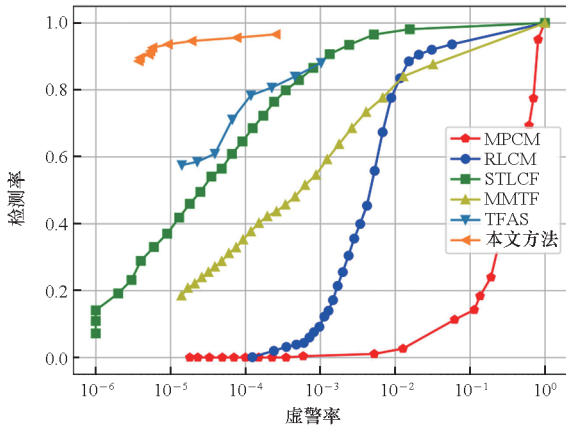
(a) 测试序列 1 的 ROC 对比曲线

(a) Comparison of ROC curves of test sequence 1



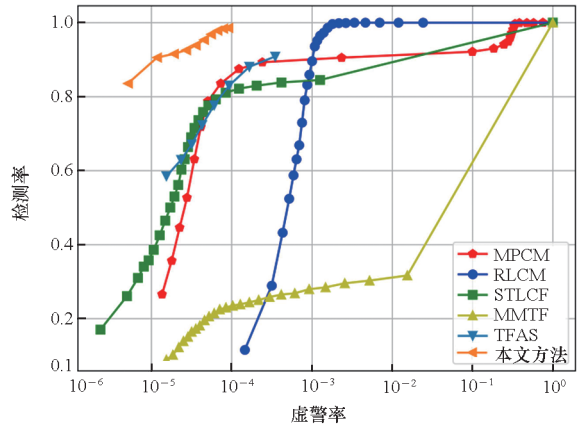
(b) 测试序列 2 的 ROC 对比曲线

(b) Comparison of ROC curves of test sequence 2



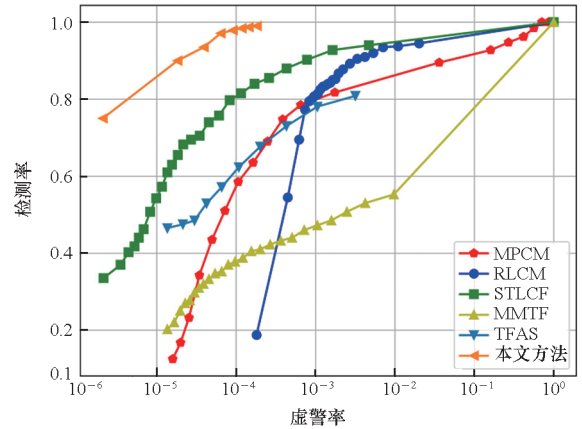
(c) 测试序列 3 的 ROC 对比曲线

(c) Comparison of ROC curves of test sequence 3



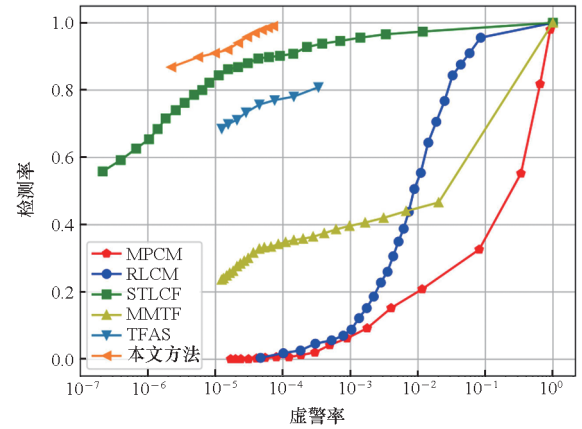
(d) 测试序列 4 的 ROC 对比曲线

(d) Comparison of ROC curves of test sequence 4



(e) 测试序列 5 的 ROC 对比曲线

(e) Comparison of ROC curves of test sequence 5



(f) 测试序列 6 的 ROC 对比曲线

(f) Comparison of ROC curves of test sequence 6

图 14 ROC 曲线比较

Fig. 14 Comparison of ROC curves

2.0 GHz, 内存为 16 GB, 操作系统为 Windows 10) 下运行,本文方法分别给出了在 X86 和 FPGA 平台下运行的平均速率。可以直观看出经过并行流水优化设计,算法在 FPGA 平台相对 X86 平台取得了显著的加速效果,充分满足实时处理应用需求。

表 4 各测试序列中不同方法平均检测率、平均虚警率与平均速率对比

Tab. 4 Comparison of average detection rate, average false alarm rate and average rate of different methods in test sequence

方法	序列 1		序列 2		序列 3		序列 4		序列 5		序列 6		平均速率/ (帧/s)
	平均检 测率/%	平均虚 警率/%	平均检 测率/%	平均虚 警率/%	平均检 测率/%	平均虚 警率/%	平均检 测率/%	平均虚 警率/%	平均检 测率/%	平均虚 警率/%	平均检 测率/%	平均虚 警率/%	
MPCM	68.58	47.99	62.25	52.35	35.05	31.96	84.16	25.33	76.11	38.28	39.54	34.49	1.92
RLCM	69.56	4.77	45.40	4.88	39.50	4.54	81.39	4.12	81.90	4.13	40.39	5.13	0.24
STLCF	65.74	3.92	73.59	3.94	57.39	3.95	60.83	3.85	67.24	3.87	82.70	3.91	0.74
MMTF	33.26	4.19	56.15	4.02	47.84	4.09	24.73	3.95	40.12	3.93	35.99	3.98	1.58
TFAS	73.63	0.05	80.49	0.02	72.40	0.02	75.06	0.01	61.45	0.05	74.31	0.08	1.15
本文 方法	92.38	0.01	94.42	0.01	93.41	0.01	91.49	0.01	91.71	0.01	94.53	0.01	0.72 100 (FPGA 平台)

3.4 FPGA 实现资源占用与性能分析

表 5 给出了优化后的检测算法实现对 FPGA 资源使用情况, FPGA 型号为 Xilinx 公司产 XC7VX690T, 开发工具为 Vivado 2019.1。如 2.2 节所述, 由于片内块随机存储器 (block random access memory, BRAM) 和用于实现寄存器功能的触发器 (flip flop, FF) 需要缓存部分权重、输入和输出部分特征图及大量中间计算结果, BRAM 和 FF 使用率分别高达 84% 与 70%。大量 BRAM 和 FF 的使用有效保证了整体并行处理架构无阻塞流水运行, 同时缓解了布局布线难度提高系统工作时钟频率。算法中乘法运算全部使用 DSP 硬核资源实现, 由于 PE 阵列较大使用率达 76%。

表 5 FPGA 资源使用情况
Tab. 5 FPGA resource utilization

资源	使用资源 数目	可用资源 数目	资源使用 率/%
BRAM_18K	2 482	2 940	84
DSP	2 748	3 600	76
FF	610 689	866 400	70
LUT	356 926	433 200	82

图 15 给出了 FPGA 仿真波形, 仿真显示单帧图像流水线上各模块最大处理时钟数为 1 442 896 个时钟周期, FPGA 布局布线后给出的最高工作频率为 167 MHz, 对应最小处理周期为 8.64 ms。实际应用时板上时钟为 150 MHz, 此时最小处理周期为 9.62 ms, 相对图像帧传输末端时延为 1.8 ms, 可有效满足 100 Hz 图像快速实时

处理需求。

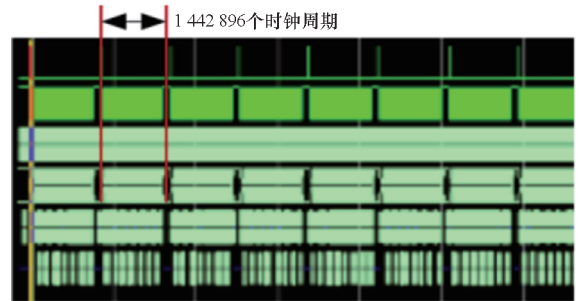


图 15 算法 FPGA 实现仿真波形

Fig. 15 Algorithm simulation waveform based on FPGA

4 结论

针对低空复杂场景下弱小目标检测难题, 本文提出一种时空域结合的目标检测方法和基于 FPGA 平台的实时并行优化处理方法。算法采用轻量化全卷积回归网络进行空域检测, 继而对空域检测结果进行连通标记和时域关联获得目标精确检测结果。测试结果表明, 在 6 个复杂低空场景弱小目标测试序列中, 本文研究成果在 0.000 1 虚警率下实现 93% 平均检测率, 较传统方法性能提升显著。并在单片 FPGA 上完成 100 Hz 图像实时处理, 处理时延低于 1.8 ms, 性能充分满足对高帧频红外图像中弱小目标快速精准实时检测应用需求。

参考文献 (References)

- [1] LI L, LI H, LI T, et al. Infrared small target detection in compressive domain[J]. Electronics Letters, 2014, 50(7): 510-512.
- [2] LIU H K, ZHANG L, HUANG H. Small target detection in

- infrared videos based on spatio-temporal tensor model [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(12): 8689–8700.
- [3] LUO K P. Space-based infrared sensor scheduling with high uncertainty: issues and challenges [J]. *Systems Engineering*, 2015, 18(1): 102–113.
- [4] QIN H L, HAN J J, YAN X, et al. Infrared small moving target detection using sparse representation-based image decomposition [J]. *Infrared Physics & Technology*, 2016, 76: 148–156.
- [5] ZHANG H, ZHANG L, YUAN D, et al. Infrared small target detection based on local intensity and gradient properties [J]. *Infrared Physics & Technology*, 2018, 89: 88–96.
- [6] ZHOU J J, LYU H Y, ZHOU F G. Infrared small target enhancement by using sequential top-hat filters [C]// *Proceedings of International Symposium on Optoelectronic Technology and Application: Image Processing and Pattern Recognition*, 2014.
- [7] 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's International Symposium on Optical Science, Engineering, and Instrumentation*, 1999: 74–83.
- [8] 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.
- [9] HAN J H, MA Y, ZHOU B, et al. A robust infrared small target detection algorithm based on human visual system [J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(12): 2168–2172.
- [10] HAN J H, LIANG K, ZHOU B, et al. Infrared small target detection utilizing the multiscale relative local contrast measure [J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(4): 612–616.
- [11] WEI Y T, YOU X G, LI H. Multiscale patch-based contrast measure for small infrared target detection [J]. *Pattern Recognition*, 2016, 58: 216–226.
- [12] DENG L Z, ZHU H, TAO C, et al. Infrared moving point target detection based on spatial-temporal local contrast filter [J]. *Infrared Physics & Technology*, 2016, 76: 168–173.
- [13] GAO J L, WEN C L, BAO Z J, et al. Detecting slowly moving infrared targets using temporal filtering and association strategy [J]. *Frontiers of Information Technology & Electronic Engineering*, 2016, 17(11): 1176–1185.
- [14] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779–788.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]// *Proceedings of European Conference on Computer Vision*, 2016: 21–37.
- [17] GULER R A, TRIGEORGIS G, ANTONAKOS E, et al. DenseReg: fully convolutional dense shape regression in-the-wild [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 6799–6808.
- [18] 回丙伟, 宋志勇, 范红旗, 等. 地/空背景下红外图像弱小飞机目标检测跟踪数据集 [J]. *中国科学数据*, 2020, 5(3): 291–302.
- HUI B W, SONG Z Y, FAN H Q, et al. A dataset for infrared detection and tracking of dim-small aircraft targets under ground/air background [J]. *China Scientific Data*, 2020, 5(3): 291–302. (in Chinese)
- [19] PANG D D, SHAN T, LI W, et al. Infrared dim and small target detection based on greedy bilateral factorization in image sequences [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 3394–3408.