**基于偏最小二乘分类器的动态视觉序列运动目标检测**

**摘要**

从视频序列中检测运动目标对目标跟踪起着至关重要的作用。本文的主要目的是利用不同的分类算法对不同的视频序列进行检测和分类。从自己拍摄的实际数据集中收集输入的视频序列，并提取单独的帧。这些帧经过预处理，然后应用于新的背景去除过程。提取了基于局部二值模式(LBP)和灰度系数的重要特征。最后利用支持向量机（SVM）、偏最小二乘法（PLS）和概率神经网络（PNN）对这些特征进行分类，并对这些分类器的性能进行了评价和比较。结果表明，PLS分类器的分类精度较高，但计算时间较长。

**关键词：**自适应直方图均衡化；局部二元模式；PLS；PNN；SVM

**1、研究背景**

视频序列中运动目标的检测和分类对目标跟踪、运动识别和视频监控具有重要意义。任何运动检测技术的目的都是将视频中运动物体的前景区域从背景区域中分离出来。常用方法采用光流技术实时检测运动目标。结果表明，该方法虽然具有较高的精度，但是同时也具有较高时间复杂度。而通过背景去除方法和时间差分方法，可以克服这些缺点和局限性。背景去除时使用输入的视频序列，通过找出帧当前像素与背景参考帧像素的差来检测帧内的运动对象。通常选择第一帧作为第一参考帧，然后按固定时间更新。而时间差分则是对视频连续帧中像素特征的差分。

光流法将图像序列赋给每个像素一个二维速度矢量。根据速度矢量的属性对运动目标进行分割，检测运动目标的边缘形状。

该方法利用背景差法检测运动目标，比较了不同分类器的性能。本文解决了应用背景去除技术所产生的许多问题。这些问题包括:光照条件的突变或渐变、场景中多个物体的移动、由于视觉源质量差或低质量而产生的帧内噪声以及前景物体作为移动物体投影的阴影区域。图1显示了背景参考帧和具有运动物体的视频帧的示例。



（a）白天静态环境



（b）夜晚静态环境



（c）白天动态环境



（d）夜晚动态环境

图1 白天和夜晚静态以及动态的视频环境

本工作的主要目标是检测视频中的运动目标，并使用不同的分类算法对其进行分类。论文的组织如下。在第2节中概述对一些主流的运动目标检测分类方法。本文提出的目标检测和分类方法将在第3节、第4节的报告中讨论，并对于本文提出方法进行的实验结果与验证。最后总结解决方案，第五部分给出结论。

**2、算法概述**

一些研究探讨了视频序列中运动目标的自动检测。Y Yang等人在应用时空模型分割和检测视频序列中移动对象的存在。从视频获取的背景帧中提取时间图像特征。为了对运动目标进行分割，采用了动态背景算法，召回率为72%，误差为27.73%以及27.04%。

Li等人提出了一种基于codebook算法的运动目标检测方法。首先，使用codebook算法对前景和背景进行分割。采用三帧差分法检测前景帧。为了消除三帧差带来的空洞，作者采用了日志边缘检测和组件填充的方法来优化前景。最后，应用逻辑‘和’运算将改进的三帧差分算法得到的前景对象与codebook算法进行关联。

Wren等人开发了一种人员分割、跟踪和解释系统。作者将该系统命名为pfinder，并对背景中的每个像素进行简单的高斯分布建模。但是利用每个像素的期望极大值估计高斯参数是一个复杂的计算问题。

Bilodeau等人提出了一种检测运动目标的有效方法。他们应用改进的局部二值相似模式从输入视频序列中分割背景。他们还测试了他们提出的方法在各种实时视频序列上的有效性。

Elgammal等人引入了一种非参数背景建模。该方法根据每个像素的灰度样本来估计观察一个像素灰度级的概率。作者利用颜色信息来抑制目标物体的阴影。

Zhang等人提出了一种视频序列中运动目标检测的训练自由方法。对于序列的每一帧，测量其自身与前一帧之间的密集光流。针对光流强的区域，提出了一种新的聚类方法。这有助于分割不同的运动物体。该方法的召回率为87.2%，准确率为93.5%。

Wu等人提出了一种利用比值图像作为运动检测基本步骤的方法。光照不足的影响被消除了。方法的选择问题通常与差值图像有关。在这种方法中，它被平移到比率图像。该问题由作者基于直方图技术自动解决。

Lanasi等人提出了一种用于运动目标检测与跟踪的非参数核密度估计方法。提出了一种基于均值漂移跟踪的多分辨率递归密度估计快速鲁棒算法。

D. Kollar等人提出了一种基于天气和时间对强度值的影响，利用卡尔曼滤波对背景像素进行建模的模型。而S Jabri等人使用颜色和边缘信息来建模背景和进行减法。作者使用置信度图来融合中间结果。当场景中出现突发变化或多个运动物体时，该方法不能产生良好的效果。

**3、提出方法**

该方法的目的是通过不同的分类算法对输入视频序列进行检测和分类。从数据集中收集所需的输入视频序列。首先将输入的视频序列转换成帧，然后对其进行预处理，以提高帧的质量，消除噪声。然后将增强后的帧应用于背景去除过程。然后提取特征向量。通过三种不同的分类器对不同视频序列的特征进行训练和测试。该方法的总体示意图如图2所示。



图2 方法总体示意图

**3.1、实验数据**

为了比较提出的工作的性能，处理所需的输入视频序列均使用实际拍摄的水下视频作为数据集。数据集中白天以及夜间的视频拍摄信息如下。

（1）白天视频

拍摄时间：2018年11月28日 13.00-14:30。

地点：上海海洋大学二教后湖里。

水深：约1.5米。

水色清澈。

天气，晴。

本次拍摄将灯关掉，没有使用灯光。

（2）夜间视频

拍摄时间：2018年11月28日 16.30-19:30。

地点：上海海洋大学二教后湖里。

水深：约1.5米。

水色清澈。

天气，晴。

本次拍摄将白灯打开作为照明。

该数据集包含不同的视频序列，在不同的环境位置由高分辨率摄像机捕获。本数据库中输入的视频序列包括白天环境条件和夜晚环境条件。通过不同的视频序列以及白天夜晚混合的视频序列来验证和比较本文提出的工作的性能。

**3.2、数据预处理**

从数据集中获得的输入视频序列首先被转换成帧。然后对每一帧进行预处理，以提高帧的质量。预处理的主要目的是通过去除噪声来提高工作的准确性。从视频中获得的每一帧都是RGB格式的。将RGB格式的帧转换成HSI（H：色调Hue，S：饱和度Saturation，I：亮度Intensity）格式，并单独提取亮度（I）部分进行进一步处理。转换公式如下。

 其中 （1）

 （2）

 （3）

提取亮度（I）部分的原因，主要是与色调以及饱和度参数相比，噪声只会对亮度产生较大的影响。然后将提取出的亮度部分应用到中值滤波器中去除噪声。

为了提高对比度，中值滤波器的输出采用对比度有限的自适应直方图均衡(CLAHE)算法。

**3.3、背景去除**

下一个重要的步骤是从背景中提取前景对象。这个背景提取算法能够回答第1节中讨论的一些关键问题。在本工作中，背景去除采用贝叶斯方法，对视频帧进行前景和背景的图像分割，前景区域的提取是通过时间图像分析和背景去除处理相结合来完成的。

首先，时间分析是通过比较两个连续的帧的时间t来执行的。设vt为视频序列中的t时刻，s=(s,y)位置为该像素的特征向量。由贝叶斯规则，vt后验概率为

 （4）

其中，式中的分母

 （5）

其中C为前景或者背景，根据贝叶斯策论，对比后验概率以及的大小，如果满足

 （6）

则把vt分类为背景，将（4）、（5）带入（6）中得到

 （7）

根据当前像素的三个概率，，来确定该像素是否为背景。

将该像素用于背景去除处理。在背景去除过程中，每次t时将图像与参考背景图像进行比较。为了检测前景对象，计算辐射相似度(RS)，公式如下。

 （8）

其中m[W]和v[W]分别表示窗口W中像素强度的均值和方差。

**3.4、特征提取**

特征提取过程的目标是用一些对分类过程有用的可量化信息来表示一个像素。在本工作中，本文选择了以下特征向量集。

(1)使用LBP的纹理特征:使用局部二值模式(LBP)算法提取基于纹理的特征。

(2)灰度特征:提取基于前景对象灰度的5个不同特征。

**3.4.1、使用LBP的纹理特征**

在文献研究的基础上，利用局部二值模式(LBP)提取了24个基于纹理的特征。LBP是用于图像处理和机器学习的强大特征描述符之一。与其他纹理描述符相比，LBP的计算复杂度非常低。

该算法的关键是在获得的前景区域中为每个像素放置一个标签。这是通过计算像素的局部邻域内点P和半径r的数量得到的。计算中心像素的强度值，并选择该值作为参考。根据这个参考值，邻域像素是形成二值模式的阈值。最后，通过将每个像素的二进制模式相加，并以2为幂进行加权，计算出LBP标签。

 （9）

 （10）

其中Ip和Ic分别为邻域像素和中心像素的灰度值，P为半径为r的圆上的样本数。

从每个LBP模式计算出六个统计特征，如均值、标准差、中位数、熵、偏度和峰度。对r =1、2、3、4的4个不同的半径进行此步骤，共得到24个特征。利用LBP处理后的图像如下图所示。

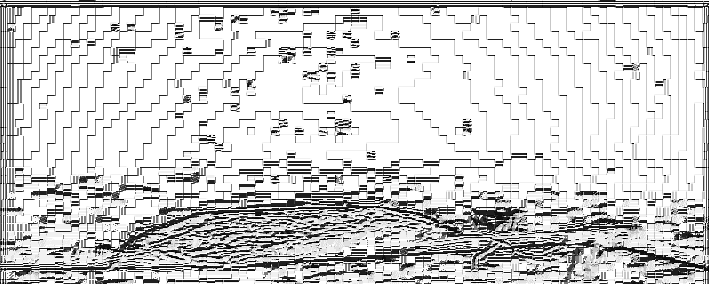


图3 利用LBP处理后的图像

**3.4.2、灰度值特征**

背景对象的灰度为输入序列的分类提供了更有意义的特征。考虑到这一信息，从前景对象中提取了一组灰度特征。设Sx,y为w x w窗口在所述像素(x,y)上运行的坐标集。这些特征如下所示。

 （11）

 （12）

 （13）

 （14）

 （15）

灰度处理后的图像如下图所示。



图4 灰度处理后的图像

**3.5、场景分类**

将提取的LBP和灰度特征组合成特征向量。为了将输入的视频序列分类为不同的类别，将这些特征向量应用到分类器算法中。本文选择了支持向量机（SVM）、偏最小二乘法（PLS）和概率神经网络（PNN）三种分类器，并对它们的性能进行了比较。

**3.5.1、SVM分类器**

将从输入视频序列的不同运动目标中提取的特征向量应用于支持向量机分类器。该分类器试图最小化经验风险，防止过度拟合问题。该分类器的结构如图5所示。

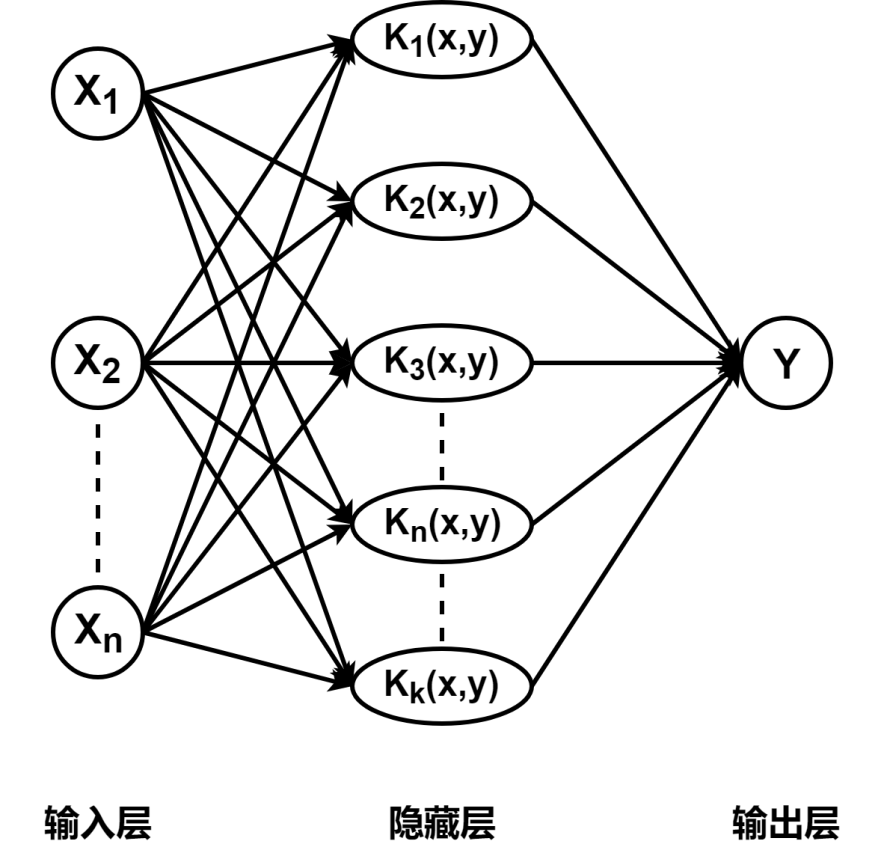


图5 SVM分类器体系结构

该分类器由输入层、隐藏层和外层三层组成。分类是在两个不同的阶段进行的（a）训练阶段（b）测试阶段。

在训练阶段，利用SVM分类器对视频帧中的白天视频和夜晚视频等不同序列提取的特征进行训练。提取的特征中有近60%用于此过程。然后，在测试阶段，对分类过程应用和测试其余的特性。该过程中使用的径向基函数核公式如下。

 （16）

其中是一个核参数。

**3.5.2、PLS分类器**

帮助对输入序列进行分类的第二种分类器是偏最小二乘法(PLS)分类器。

该分类器在不同分类之间具有低偏差和高方差。本文采用一种阈值可调的线性回归PLS分类器。选择该分类器的主要原因是它提供了较高的精度，避免了过拟合问题。通常，这个分类器被表示为

 （17）

其中，A为具有分类度量的向量，B为提取的特征向量。β是线性回归系数和ε是一个残余向量。

将提取的特征向量B（见3.4节）应用于PLS分类器进行训练，得到最优的线性回归系数。将此最优值应用于测试阶段，对输入序列进行分类。

**3.5.3、PNN分类器**

接下来对概率神经网络分类器进行了分类实验。该分类器是多层前馈神经网络分类器的一种，由贝叶斯网络派生而来。该分类器由输入层、模式层、求和层和输出层四层组成，如图6所示。

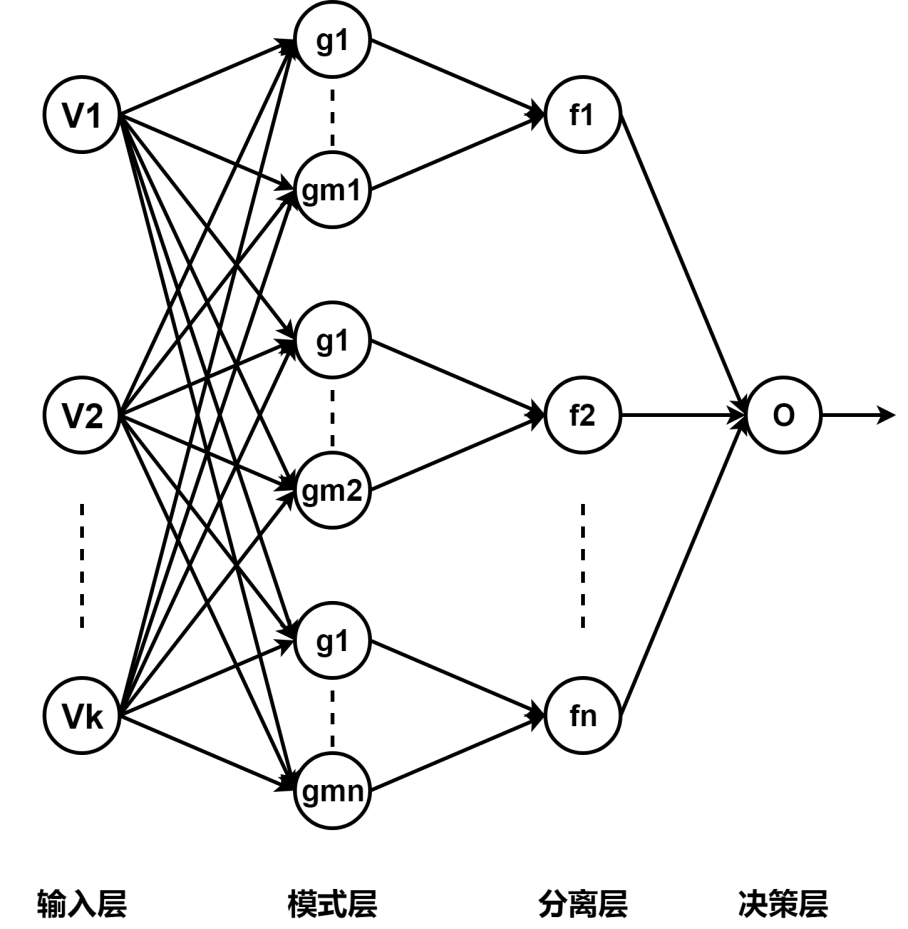


图6 PNN分类器体系结构

该PNN的步骤过程如下:

步骤1:将提取的输入特征输入到输入层。

步骤2:将训练好的特征输入到下一个模式层。公式如下所示。

 （18）

其中，Xki是核的中心部分，σ是平滑参数。这个σ决定了核函数的运行效率。

步骤3:下一步是比较每组模式的输入特性。利用这个比较，求和层利用先前计算的密度的组合来计算条件概率，公式如下。

 （19）

其中Wki表示满足的正系数，。

步骤4:在每个类输出节点上，将具有与输入相似特性的所有帧添加到输入节点。

步骤5:使用以下公式计算输出节点上所有增加的函数值的最大值。公式如下。

 （20）

输出层给出不同的分类结果。

**4、实验结果与讨论**

该方法的性能已经在水下环境视频序列上进行了测试。具体选择水下环境拍摄的白天视频和夜晚视屏进行验证和比较。考虑了不同场景条件下的视频帧序列。

该算法是在Intel(R) Core(TM) i5-8500 CPU @ 3.00 GHz处理器上使用MATLAB (R2013a)运行的。通过计算精度、检测率（DR）、误检率（FAR）和计算时间，获得了本文工作的定量性能。

 （21）

 （22）

 （23）

其中，TP为准确率（True Positive），FP为误报率（False Positive），FN为漏报率（False Negative）准确率表示检测到的与移动物体相对应的像素的数量。误报率表示检测到的不符合移动对象的像素数量，漏报率表示未检测到的移动对象像素。

在表1中，所提出的方法与其他方法的表现比较。表1中提出的工作的精度是使用PLS分类器获得的不同视频序列的平均精度。从表中可以看出，与文献中讨论的其他方法相比，该方法具有较高的准确性。

表1 使用PLS分类器提出的方法与其他最先进的方法的比较

|  |  |
| --- | --- |
| **方法** | **精度（%）** |
| 提出的方法 | 94.58 |
| Patel等人（2014） | 91.53 |
| Barnich等人（2011） | 89.64 |
| St. Charles等人（2014） | 87.61 |

表2显示了SVM分类器对不同序列的分类实验结果。

表2 利用SVM分类器对输入序列进行分类得到实验结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **视频序列** | **精度（%）** | **检测率（DR，%）** | **错检率（FAR，%）** | **计算时间（s）** |
| 白天视频 | 89.53 | 86.20 | 13.60 | 56.55 |
| 夜晚视频 | 87.40 | 84.41 | 14.20 | 59.43 |
| 混合视频 | 84.51 | 81.64 | 15.39 | 62.38 |
| 未处理白天视频 | 84.39 | 81.42 | 16.38 | 57.98 |
| 未处理夜晚视频 | 81.75 | 77.98 | 19.82 | 59.74 |
| 未处理混合视频 | 78.36 | 74.57 | 23.23 | 62.26 |

表3显示了PLS分类器对不同序列的分类实验结果。

表3 利用PLS分类器对输入序列进行分类得到实验结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **视频序列** | **精度（%）** | **检测率（DR，%）** | **错检率（FAR，%）** | **计算时间（s）** |
| 白天视频 | 94.58 | 92.32 | 8.20 | 65.27 |
| 夜晚视频 | 91.95 | 89.98 | 9.30 | 67.18 |
| 混合视频 | 89.62 | 87.42 | 10.38 | 69.52 |
| 未处理白天视频 | 88.53 | 85.37 | 12.43 | 66.78 |
| 未处理夜晚视频 | 85.91 | 82.92 | 14.88 | 68.65 |
| 未处理混合视频 | 83.49 | 79.43 | 18.37 | 70.84 |

表4显示了PNN分类器对不同序列的分类实验结果。

表4 利用PNN分类器对输入序列进行分类得到实验结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **视频序列** | **精度（%）** | **检测率（DR，%）** | **错检率（FAR，%）** | **计算时间（s）** |
| 白天视频 | 87.62 | 86.14 | 14.22 | 53.92 |
| 夜晚视频 | 85.23 | 84.20 | 15.10 | 55.30 |
| 混合视频 | 83.53 | 82.53 | 15.27 | 57.67 |
| 未处理白天视频 | 81.10 | 79.72 | 18.08 | 54.48 |
| 未处理夜晚视频 | 78.94 | 77.81 | 19.99 | 55.76 |
| 未处理混合视频 | 75.16 | 74.48 | 23.32 | 56.19 |

从以上结果可以看出，分类器对于白天视频的分类效果最好，夜晚视频精度略有下降，而白天和夜晚混合的视频精度并不理想。从总体精度可以看出，使用PLS分类器的分类精度比较好，但总体计算时间较长。对于未经过本文方法处理过后的视频数据，分类精度普遍比经过本文方法处理后的视频数据低，但计算时间差异不大。

图7显示了三种不同分类器的性能比较。在这个比较过程中，对比了不同分类器的精度和计算时间。

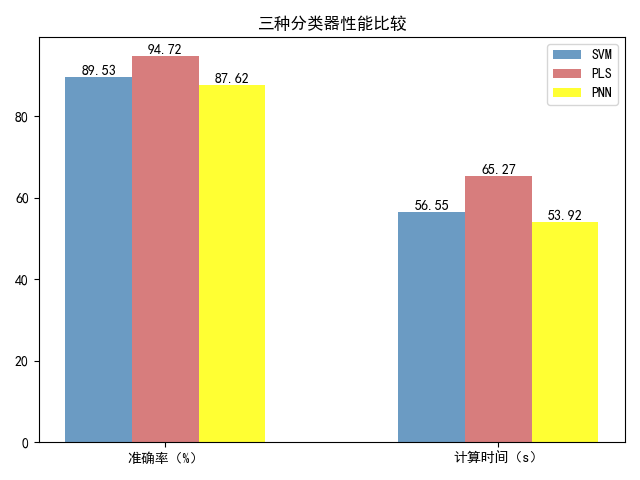


图7 三种分类器性能比较

**5、结论**

在这篇论文中，比较分析了不同的分类器对运动目标的检测。这项工作在目标跟踪、视频监控、运动识别等方面非常有用。提出了一种视频序列中运动目标检测的鲁棒算法。该方法首先提取输入视频，然后提取图像序列。这些序列是预处理和前景对象单独分割的背景去除技术。利用三种不同的分类器，提取了基于LBP和灰度系数等不同有用信息，并对不同场景进行了分类。

三种分类器在白天和夜晚以及混合环境下的结果表明了该方法的鲁棒性和可靠性。本文的缺点是不能完全消除场景中由于背景的高对比度所造成的的阴影。

**6、参考文献**

[1] Barnich, and Van Droogenbroeck, M., ViBe: A universal background subtraction algorithm for video sequences. IEEE Trans.Image Process. 20(6):1709–1724, 2011.

[2] Patel, C. I., Garg, S., Zaveri, T., and Banerjee, A., Top-down and bottom-up cues based moving object detection for varied background video sequences. Advances in Multimedia 2014:13, 2014.

[3] Yang, Y., Zhang, Q., Wang, P., Hu, X., and Wu, N.: 'Moving object detection for dynamic background scenes based on spatiotemporal

model', Advances in Multimedia, 2017.

[4] Li, S., Liu, P. and Han, G.: 'Moving object detection based on codebook algorithm and threeframe difference', International Journal of Signal Processing, Image Processing and Pattern Recognition, 2017, 10(3), pp. 23-32.

[5] Wren, C., Azarbayejani, A., Darrell, T., and Pentland, A., Pfinder:Real-time tracking of the human body. IEEE Trans. Pattern Anal. Mach. Intell. 19(7):780–785, 1997.

[6] Bilodeau, G. A., Jodoin, J. P., and Saunier, N., 'Change detection in feature space using local binary similarity patterns,' in Proceedings of the international conference on Computer & Robot Vision, 10(1), pp. 106–112, , 2013.

[7] Elgammal, A., Harwood, D., Davis, L.:, 'Non-parametric model for background subtraction' IEEE frame rate workshop, 1999, 21 September, Corfu, Greece.

[8] Zhang, Y., Li, G., Xie, X. and Wang, Z.:, 'A new algorithm for fast and accurate moving object detection based on motion segmentation by clustering', 2017 fifteenth IAPR international conference on machine vision applications (MVA), Nagoya, pp. 444-447, , 2017.

[9] Wu, Q. Z., Cheng, H. Y., and Jeng, B. S., Motion detection via change-point detection for cumulative histograms of ratio images. Pattern Recogn. Lett. 26(5):555–563, 2005.

[10] Ianasi, C., Gui, V., Toma, C. I., and Pescaru, D., A fast algorithm for background tracking in video surveillance, using nonparametric kernel density estimation. FactaUniversitatis Ser.: Elect. Energy 18(1):127–144, 2005.

[11] Koller, D., Weber, J., Huang, T., Malik, J., Ogasawara, G., Rao, B., Russel, S.: 'Towards robust automatic traffic scene analysis in realtime' Proceedings of the international conference on pattern recognition, Israel, 1994.

[12] Jabri, S., Duric, Z., Wechsler, H., Rosenfeld, A.: ' Detection and location of people in video images using adaptive fusion of colour and edge information', ICPR 2000, 3–8 September, Barcellona, Spain, pp. 627–630.

[13] Ramasubramanian, B., and Selvaperumal, S., A novel efficient approach for the screening of new abnormal blood vessels in color fundus images. Appl. Mech. Mater. 573:808–813, 2014.

[14] Lu, X. et al: ,'Feature extraction and fusion using deep convolutional neural networks for face detection', Hindawi Mathematical Problems in Engineering, 2017.

[15] Marín, D., Aquino, A., Gegúndez-Arias, M. E., and Bravo, J. M., A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. IEEE Trans. Med. Imaging 30(1):146–158, 2011.

[16] Ramasubramanian, B., An efficient integrated approach for the detection of exudates and diabetic maculopathy in colour fundus images. Advanced Computing: An International Journal (ACIJ) 3(5): 83–91, 2012.

[17] Agurto, C. et al., A multiscale optimization approach to detect exudates in the macula. IEEE Journal of Biomedical and Health Informatics 18(4):1328–1336, 2014.

[18] de Jong, S., SIMPLS: An alternative approach to partial least squares regression. ChemometricsIntell. Lab. Syst. 18:251–263, 1993.

[19] Ramasubramanian, B., Selvaperumal, S. : 'Efficient approach for the automatic detection of haemorrhages in colour retinal images', IET Image Processing, Online ISSN 1751-966.

[20] Vl, G., Pavlidis, N. G., Parsopoulos, K. E. et al., New self-adaptive probabilistic neural network in bio-informatics and medical tasks. International Journal on Artificial Intelligence Tool 15(3):371–396, 2006