# Fire Detection Using Level Set Segmentation Based Fractional Order Optical Flow and 4D Fire Features with Mixed Data CNN-LSTM Model

Muzammil Khan and Pushpendra Kumar
Department of Mathematics, Bioinformatics and Computer Applications
Maulana Azad National Institute of Technology Bhopal-462003, India
{mk.193104002,pkumarfma}@manit.ac.in

***Abstract*—As we are aware that the world witnesses a huge number of fire breakouts everyday, which results in high numbers of hazardous events and severe losses to property and forest vegetation. Therefore, early stage fire detection is of vital importance, for once it spreads it becomes unmanageable and disastrous. The early detection of fire can be performed with the help of vision based deep learning techniques. The novelty of the work lies in performing the fire detection using the static and dynamic features of fire. The static fire features are taken as shape, texture, and color, while the dynamic feature accounts for its flickering motion. For this purpose, the fire motion is estimated in terms of optical flow from videos (image sequences) by using a motion edge preserving level set segmentation based fractional order variational model. Level sets provide nicely segmented flow fields, while fractional order derivatives are capable to deal with discontinuities in the motion field. The estimated optical flow field is used to derive four fire features, which are constituted as 4D vectors. These 4D vectors reduce the data dimensionality and mitigates over-fitting problem. Finally, the fire detection is carried out by implementing a mixed data CNN-LSTM model. The mixed data presented in the work is composed of a reference image frame and the corresponding 4D vector sequence. Also, the significance of the model is manifested through an ablation study. The model performance validation is performed thorough a comparison study conducted with several existing models.**

**Keywords**- CNN-LSTM, Fire detection, Fractional order optical flow, Level set segmentation, Static and dynamic features.

## I. Introduction

As we know that the conventional fire detectors rely on temperature and smoke [1], which makes them highly sensitive to the presence of humidity, dust, pollen, cooking fumes, environment temperature fluctuations, and system sparking, etc. [2]. This results in false fire alarms. Also, these detectors exhibit limited applicability when implemented in larger areas and for longer distances.

In recent years, image-based fire detection utilizing computer vision algorithms has garnered significant attention as an alternative solution [3]. Xu et al. [4] introduced an ensemble learning based framework using a dual learner Yolov5-EfficientDet and an individual learner EfficientNet for identifying forest fires. Li et al. [5] developed a deep learning model composed of three multi-scale feature extraction modules and four implicit deep supervision modules with channel attention in fire detection. An et al. [6], implemented a dynamic convolution YOLOv5 fire detection method and performed anchor box optimization by utilizing K-mean++ algorithm to reduce the error rate of the method. All such studies are based on image features such as color, texture and shape. Many researchers have demonstrated that several distinguishing features of fire can be obtained from its flickering motion in an image sequence [7]. This makes the extraction of fire motion content in an image sequences extremely crucial, which is generally represented in terms of optical flow.

Optical flow is a $2D$ vector field which establishes a correspondence between the pixels of two consecutive image frames in an image sequence [8]. Computation of optical flow is generally done by using variational approaches. These approaches encode the optical flow properties in the form of an energy functional. These properties generally include flow field smoothness, preservation of motion edges, and denseness. The minimization of the resulting energy functional provides a system of equations for optical flow estimation [8], [9]. However, these are integer order derivative models, and unable to illustrate the flickering motion of fire accurately. The works presented in [10]–[13] provide variational model generalization from integer to fractional order and demonstrated an increased accuracy in optical flow applications. The fractional order derivatives are capable of producing better results compared to integer order derivative for discontinuous functions such as images [14]. Further, Khan et al. [15] exhibited the significance of level set segmentation (LSS) framework in fractional order variational techniques and shown nicely segmented optical flow fields. This work is aimed at utilizing fractional order optical flow for performing fire detection.

This paper proposes a novel binary classification framework for fire detection in videos based on static and dynamic features of fire. Static features are incorporated as shape, texture, and color of fire, while the dynamic features are derived from fire motion. The fire motion is estimated in terms of optical flow using an LSS based fractional order variational model. The estimated optical flow field is employed in calculating four distinct features associated with fire motion. These features are organised in the form of 4D vectors, which assists in data dimensionality reduction. The reference images and their corresponding 4D vector

sequences are used to produce mixed data. This data is employed in training and testing of the proposed mixed data CNN-LSTM (convolutional neural network-long short-term memory) model. The motivation behind utilizing a CNN-LSTM model instead of a CNN model is to fuse the convolution based processing capabilities of CNN with time-series dealing strength of LSTM. Moreover, LSTM has a hierarchal structure that makes it a potential candidate for performing complex feature extraction in comparison to traditional machine learning approaches [16]. Also, an ablation study is conducted to discuss the significance of each component of the presented framework. The results are compared with several state-of-the-art models to validate the performance of the propounded research.

The forthcoming part of the research paper is outlined as: Section II outlines the methodology, Section III delves into the results and discussions of the experiments, and Section IV wraps up the paper with conclusion.

## II. Methodology

The dataset utilized in the presented work is taken from [17], which contains different image sequences extracted from several fire and non-fire videos. The total number of image frames for fire and non-fire are $27,461$ and $11,392$, respectively. These image sequences correspond to different scenes such as forest, hotel, market, traffic, sports ground, crowds, illumination change, etc. From these image sequences, $35$ and $43$ image sequences are randomly chosen for experiments in fire and non-fire classes, respectively. Thus, the prepared dataset contains total $3,603$ and $3,349$ image frames for fire and non-fire videos, respectively. The motivation behind selecting only a few image frames from the entire dataset [17] is that each video contains a sequence of image frames from a single scene. Therefore, image frames in a video are correlated with one-another. Thus, extracting and utilizing all the image frames from a video make training and testing computationally less efficient and lead to overfitting. In order to deal with these issues, only a few image frames are taken instead of the entire dataset [17]. The size of each frame is $128 \times 128$. Further, the dataset is preprocessed in accordance with [7] to highlight the fire content in the image frames. These preprocessed dataset is passed to the LSS based fractional order variational model to generate the sequences of optical flow fields for fire motion. Also, the preprocessed image sequences are thresholded to produce binary masks. Now, utilizing these binary masks with their corresponding optical flow sequences, the $4D$ feature vectors are calculated from the localized fire region. Finally, by the fusion of first image frame from an input image sequence and its corresponding sequence of $4D$ feature vectors, a mixed dataset is obtained. This dataset is employed in training and testing of a novel CNN-LSTM model. The dataset is split into a ratio 8:2 for training and testing purpose. Hence, the numbers of training and testing images are $5,561$ and $1,390$, respectively. The overall methodology to the proposed framework is demonstrated in Fig. 1. Each step of the methodology is described in detail in the forthcoming sections.

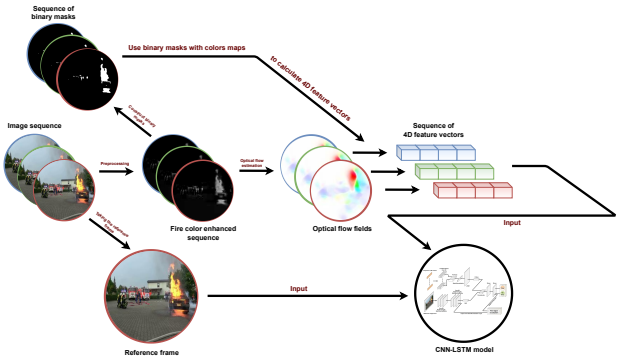


Fig. 1. Methodology for the proposed work.

### A. Fire color region enhancement

In general, a scene may contain several objects that are moving over time. Therefore, for an accurate estimation of fire motion, it becomes important to suppress the background object motion. For this purpose, Mueller et al. [7] emphasized the importance of the Hue term in the HSV color space. They found that the objects like fire exhibit a Hue value in the vicinity of $0.083$ and $1.0$. In the presented work, to highlight the fire color in each frame of an image sequence, the RGB color space is transformed into HSV color space. The values of hue component which lie near $0.083$ and $1.0$ are retained, while the remaining values are smothered by using the logistic function as given in Mueller et al. [7]

$$\mathcal{F}_l(z) = \frac{e^{-a.(z-b)}}{1 + e^{-a.(z-b)}} \tag{1}$$

where, $z$ is any variable, $a$ and $b$ are the parameters, those are chosen to weigh the fire flames in the image. Using this function, fire highlighted images are obtained as grayscale images

$$\mathcal{I} = \mathcal{F}_l(\min\{|0.083 - H|, 1 - |0.083 - H|\}) \odot S \odot V \tag{2}$$

here, H, S, and V are the hue, saturation and intensity values of the image, respectively and $\odot$ is the Hadamard product of matrices. The complete process is described in Fig. 2.

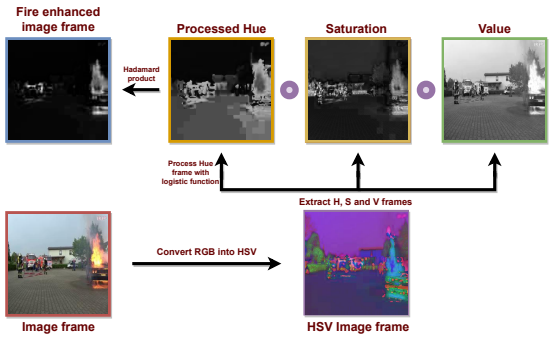


Fig. 2. Producing fire enhanced image frames.

### B. Fire optical flow estimation using LSS based fractional order variational model

The proposed model is described by the following expression [15]

$$\mathcal{F} = \lambda \mathcal{F}_{dp} + \mathcal{F}_{sc} \tag{3}$$

where, $\mathcal{F}_{dp}$ and $\mathcal{F}_{sc}$ are the segmented data penalty and smoothness terms, respectively. The forthcoming subsections depict the mathematical formulation of $\mathcal{F}_{dp}$ and $\mathcal{F}_{sc}$.

*1) Segmented data penalty term:* Let $\pi_0$ and $\pi_0^{\mathrm{C}}$ represent two disjoint image regions. Thus, the desired penalty term is implemented with Charbonnier norm as follows [15]

$$\mathcal{F}_{dp} = \int_{\pi_0} \sqrt{\kappa^2 + (\mathcal{I}_t + \nabla \mathcal{I}^T \boldsymbol{w})^2} d\boldsymbol{X} + \int_{\pi_0^{\mathrm{C}}} \sqrt{\kappa^2 + (\mathcal{I}_t + \nabla \mathcal{I}^T \boldsymbol{w})^2} d\boldsymbol{X} \tag{4}$$

where, $\kappa$ is a Charbonnier norm parameter [18], $\nabla \mathcal{I} = (\mathcal{I}_x, \mathcal{I}_y)^T$, and $\mathcal{I}_x$, $\mathcal{I}_y$, and $\mathcal{I}_t$ denote the partial differentials of $\mathcal{I}$ with respect to $x$, $y$, and $t$, respectively. Here, $\boldsymbol{w} = (u, v)^T$ is the optical flow, $\pi_0$ describes a component of the region segmented by $\mathcal{L}_\cup = \mathcal{L}_u \cup \mathcal{L}_v$, while $\pi_0^{\mathrm{C}}$ is the complements of $\pi_0$. Here, $\mathcal{L}_u$ and $\mathcal{L}_v$ are the $u$ and $v$-component level sets [19], respectively. This expression (4) is further simplified by following the definition of Heaviside's function $\mathcal{H}$ as

$$\begin{aligned} \mathcal{F}_{dp} = \int_{\mathcal{R}} \Big[ & \mathcal{H}(\Lambda_\cup) \sqrt{\kappa^2 + (\mathcal{I}_t + \boldsymbol{w}^T \nabla \mathcal{I})^2} \\ & + (1 - \mathcal{H}(\Lambda_\cup)) \sqrt{\kappa^2 + (\mathcal{I}_t + \boldsymbol{w}^T \nabla \mathcal{I})^2} \Big] d\boldsymbol{X} \end{aligned} \tag{5}$$

here, $\Lambda_\cup$ represents the level surface corresponding to the level set $\mathcal{L}_\cup$, and $\mathcal{R} = \pi_0 \cup \pi_0^C$. On simplification of (5) gives

$$\mathcal{F}_{dp} = \int_{\mathcal{R}} \sqrt{\kappa^2 + (\mathcal{I}_t + \boldsymbol{w}^T \nabla \mathcal{I})^2} d\boldsymbol{X} \tag{6}$$

This is desired data penalty term.

*2) Segmented smoothness constraint:* The smoothness constraint incorporated with fractional order derivatives is defined as

$$\mathcal{F}_{sc} = \int_{\mathcal{R}} \|\mathcal{D}^\alpha \boldsymbol{w}^T\|_F d\boldsymbol{X} \tag{7}$$

where $\boldsymbol{w} = (u, v)^T$, the Marchaud fractional order derivative of order $\alpha \in (0, 1)$ is denoted by $\mathcal{D}^\alpha = (\mathcal{D}_x^\alpha, \mathcal{D}_y^\alpha)^T$, and $\|.\|_F$ represents Frobenious norm. The rationale of utilizing Marchaud fractional derivative is its ability to deal with non-differentiable functions.

Thus, the LSS based smoothness term of the above expression (7) is defined as [19]

$$\mathcal{F}_{sc} = \int_{\pi_0} \|\mathcal{D}^\alpha \boldsymbol{w}^T\|_F d\boldsymbol{X} + \int_{\pi_0^{\mathrm{C}}} \|\mathcal{D}^\alpha \boldsymbol{w}^T\|_F d\boldsymbol{X} + \vartheta \int_{\mathcal{R}} \|\nabla \mathcal{H}_w^T\|_{\mathrm{Col}_2} d\boldsymbol{X} \tag{8}$$

where the parameter $\vartheta$ is defined as in [19], $\mathcal{H}_w = (\mathcal{H}(\Lambda_u), \mathcal{H}(\Lambda_v))^T$, and the symbol $\|.\|_{\mathrm{Col2}}$ stands for a norm function that computes the sums of the column of $L_2$-norm values in a matrix. The formulation of (8) allows the proposed model to locally smoothen the distinct region component corresponding to both the $u$ and $v$ components. Further simplification of (8) follows the use of Heaviside's function $\mathcal{H}_\omega$ as

$$\mathcal{F}_{sc} = \int_{\mathcal{R}} \|\mathcal{D}^\alpha \boldsymbol{w}^T \mathcal{H}_w^{d\frac{1}{2}}\|_F + \|\mathcal{D}^\alpha \boldsymbol{w}^T (I - \mathcal{H}_w^d)^{\frac{1}{2}}\|_F + \vartheta \|\nabla \mathcal{H}_w^T\|_{\mathrm{Col}_2} d\boldsymbol{X} \tag{9}$$

where, the superscript $d$, and $I$ represent the diagonal and identity matrices, respectively. Expression in (9) is solved following [15], [20] to obtain

$$\mathcal{F}_{sc} = \int_{\mathcal{R}} \|\mathcal{D}^\alpha \boldsymbol{w}^T\|_F + \vartheta \|\nabla \Lambda_w \delta_w^d\|_{\mathrm{Col}2} d\boldsymbol{X} \tag{10}$$

where, $\Lambda_w = (\Lambda_u, \Lambda_v)^T$. Here, $\delta$ denotes the dirac's delta function such that $\delta_w = (\delta(\Lambda_u), \delta(\Lambda_v))^T$. The expression in (10) performs vertically and horizontally oriented segmentation of the flow field.

Thus, the proposed fractional order variational model (3) can be written as

$$\mathcal{F} = \int_{\mathcal{R}} \left[ \lambda \sqrt{\kappa^2 + (\mathcal{I}_t + \boldsymbol{w}^T \nabla \mathcal{I})^2} + \|\mathcal{D}^\alpha \boldsymbol{w}^T\|_F + \vartheta \|\nabla \Lambda_w \delta_w^d\|_{\mathrm{Col}2} \right] d\boldsymbol{X} \tag{11}$$

Solution procedure of (11) involves its decomposition into multiple simpler functionals in accordance with [21]. The minimization of these component functionals is carried out using calculus of variation. The resulting system of Euler-Lagrange equations is as follows

$$\left( \frac{\lambda}{\mathcal{T}} \nabla \mathcal{I} \nabla \mathcal{I}^T + \frac{1}{\Phi} \mathcal{I} \right) \hat{\boldsymbol{w}} = \frac{1}{\Phi} \boldsymbol{w} - \frac{\lambda}{\mathcal{T}} \mathcal{I}_t \nabla \mathcal{I} \tag{12}$$

$$u^+ - \hat{u} = -2\Phi \{(\mathcal{D}_-^\alpha \mathcal{D}_+^\alpha)^T e\} u^+ \text{ on } \{(x, y) | \Lambda_u(x, y, t) > 0\} \tag{13}$$

$$u^- - \hat{u} = -2\Phi \{(\mathcal{D}_-^\alpha \mathcal{D}_+^\alpha)^T e\} u^- \text{ on } \{(x, y) | \Lambda_u(x, y, t) < 0\} \tag{14}$$

$$\begin{aligned} \frac{\partial}{\partial t} \Lambda_w(x, y, t) = & \\ & - \delta_w^d \Big[ \frac{1}{2\Phi} (\boldsymbol{w}^+ - \hat{\boldsymbol{w}})^d (\boldsymbol{w}^+ - \hat{\boldsymbol{w}}) - \frac{1}{2\Phi} (\boldsymbol{w}^- - \hat{\boldsymbol{w}})^d (\boldsymbol{w}^- - \hat{\boldsymbol{w}}) \\ & + diag\{(\mathcal{D}^\alpha \boldsymbol{w}^+)^T (\mathcal{D}^\alpha \boldsymbol{w}^+)\} - diag\{(\mathcal{D}^\alpha \boldsymbol{w}^-)^T (\mathcal{D}^\alpha \boldsymbol{w}^-)\} \\ & - \vartheta \left\{ \nabla^T \left\{ (\nabla \Lambda_w^T) \left( diag(diag((\nabla \Lambda_w^T)^T (\nabla \Lambda_w^T))) \right)^{-\frac{1}{2}} \right\} \right\}^T \Big] \end{aligned} \tag{15}$$

where, the parameter $\Phi$ is a small real number and provides a close approximation of $\boldsymbol{w}$ by $\hat{\boldsymbol{w}}$, and $\mathcal{T} = \sqrt{\kappa^2 + (\mathcal{I}_t + \hat{\boldsymbol{w}}^T \nabla \mathcal{I})^2}$. Here, $\partial t$ is the level set function evolution time step. Similarly, the $v$ component provides a system of equations. For numerical implementation purpose, these equations can be discretized using the theory available in [19], [21].

### C. Formation of $4D$ feature vector

Once, the fire optical flow field is estimated, it can be used to extract the four $4D$ feature vector elements as follows

1) **Mean fire flow (MFF):**
   This feature is based on the flow magnitude of the fire in a scene, which is calculated using fire localized optical flow field. This estimates the mean of fire flow magnitude. Its is defined as

$$feat_1 = \underset{\mathcal{R}_F}{\mathrm{mean}}(\|\boldsymbol{w}\|_2^2) \tag{16}$$

here, $\mathcal{R}_F$ represents the localized fire region in an image frame.

2) **Weighted mean fire flow (WMFF):**
This feature is based on weighted fire flow magnitude and estimated as the mean of the product of the localized optical flow field with its corresponding fire color enhanced image. It is given as

$$feat_2 = \underset{\mathcal{R}_F}{\text{mean}}(\mathcal{I} \odot \|\boldsymbol{w}\|_2^2) \tag{17}$$

The features $feat_1$ and $feat_2$ produce high values for fire colored objects, which also includes rigid objects. Thus, the fire can be further distinguished from these rigid objects by using the next two features that give large values for turbulent fire motion.

3) **Sink/Source matching:**
Since fire motion is curl free [7], therefore, the turbulent motion of fire results in the formation of sources and sinks in the estimated optical field, whereas rigid objects tend to represent a motion mostly composed of parallel flow vectors. Thus, in order to calculate this feature, an ideal fire flow template $\mathcal{I}_f$ is designed and implemented as a kernel with odd dimensions. The location of $(x, y) = (0, 0)$ corresponds to the center pixel in the kernel window. It is convolved with the estimated flow field $\boldsymbol{w}$. This feature is obtained as the maximum of the final convolved image. It is written as

$$\mathcal{I}_f(\boldsymbol{x}) = e^{-\|\boldsymbol{x}\|}\boldsymbol{x} \tag{18}$$

$$feat_3 = \max_{\mathcal{R}} |\mathcal{I}_f \star \boldsymbol{w}| \tag{19}$$

where, $\mathcal{R}$ is the image frame, $\boldsymbol{x} = (x, y)^T$ and $\star$ represents convolution operation

4) **Flow direction variance:**
Since, fire colored rigid objects generally move in a certain direction, that results in a large variance in the flow directions. While, fire demonstrates motion in all the directions, which causes the flow field to have a low directional variance value. The flow field is used to produce a $uv$ plot, where a flow vector is denoted as a point $(u, v)$. This plot is further broken down into $n$ wedge shaped regions $r_1$, $r_2$,..., $r_n$. Then the number of vectors $m_i$ contained in each region $r_i$ is derived and normalized by dividing it from the total number of flow vectors. This feature is calculated by using the following expression

$$feat_4 = \text{Variance}\{m_1, m_2, \ldots, m_n\} \tag{20}$$

### *D. Proposed mixed-data CNN-LSTM model*

These four features are organised to form a $4D$ vector corresponding to the input reference images. Thus, using the reference image and its corresponding sequence of $4D$ vectors, a mixed dataset is prepared as described in methodology. This mixed dataset is then employed in training and testing of the proposed CNN-LSTM model. Now, the reference images are taken as an input to the CNN branch of the model, which is composed of two conv2D layers with 16 and 32 filters, two maxpooling and batchnormalization layers, and one GAP (global average pooling) layer. This branch outputs a feature vector of length 32. On the other end of the CNN-LSTM architecture, the corresponding sequences of $4D$ vectors are passed as the input. This branch is comprised of timedistributed conv2D, reshape, batchnormalization, LSTM and flatten layers. The time step taken is 10 and the output of this branch is a vector of length 80. Further, these output vectors are concatenated to construct the output feature vector and fed into a sequence of two dense layers containing 32 and 2 units, respectively for binary classification purpose. Further, in order to localize the fire region in the reference image frame, the output feature map from the batchnormalization layer of the CNN branch is fused with the weight vector corresponding to fire node in the classifier layer. The complete architecture is shown in Fig. 5 and contains only $10,818$ parameters, which makes it simple and to be implemented even in mobile devices. The complete pseudo code of the presented framework is given in Algorithm 1.

## III. Experimental results and discussions

The dataset used in experiments is curated from [17] which contains $27,461$ fire and $11,392$ non-fire images.

First experiment discusses the classification results in terms of probability score and fire region localisation for the presence of fire in a scene. This is illustrated in Fig. 3 on some sample reference image frames (bottom left corner). In Fig. 3, first row contains fire images, while the second row corresponds to non-fire images. These images show both the indoor and outdoor scenes. The first image on the first row is mostly composed of background objects such as fire colored vehicles, trees and buildings, which makes fire detection difficult. Despite these obstacles, the fire is successfully detected and localised in this video with a probability score $0.97$. The second and third images in this row show small and saturated fire flames, respectively, moreover in the third image, the fire is in its initial stage and resembles a lamp. But, the proposed algorithm provides an accurate fire detection and localisation for both these scenes with probability scores $0.85$ and $0.99$, respectively. In the second row, the first image describes a scene of a moving orange colored truck, while the second image represents a view from the ground towards the saturated sky through trees, and the third one gives a look of a scene containing an actual electric lamp. The probability score for each of these three videos are calculated as $0.01$, $0.01$ and $0.02$, respectively. This experiment validates the applicability of the proposed algorithm under different conditions.

Second experiment manifests the performance of the proposed algorithm in terms of ROC curve as shown in Fig. 4. This figure represents a nearly rectangular curve, and has the area under the curve (AUC) score of $0.999$. This high AUC score infers a satisfactory classification performance of the proposed framework.

Third experiment establishes the proposed algorithm as a state-of-the-art algorithm for fire detection by providing a comprehensive comparison with other existing algorithms. The comparison is exhibited in Table I. In order to compare

**Algorithm 1:** Algorithm to prepare mixed data

**Input**: Image sequence $\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_n$
**Output**: Mixed data for training and testing CNN-LSTM model

**for** $r$ = 1 *to* $n$ **do**
  $H, S, V$ = convert_rgb_to_hsv($\mathcal{I}_r$)
  $\mathcal{I}_r^F$ = fire_enhanced_image($H, S, V$);
**for** $s$ = 1 *to* $n$ **do**
  Initialize $\mathcal{I}_s^{mask}$ as a zero matrix of $size(\mathcal{I}_s^F)$;
  **for** $x = 1$ *to* $end$ **do**
    **for** $y = 1$ *to* $end$ **do**
      **if** $\mathcal{I}_s^F(x,y) \geq c \times \max \mathcal{I}_s^F$ **then**
        $\mathcal{I}_s^{mask}(x,y) = 1$;
      **else**
        $\mathcal{I}_s^{mask}(x,y) = 0$;
**for** $p$ = 1 *to* $n-1$ **do**
  Set up image pyramids $\mathcal{I}_p^{pyr(k)}$ and $\mathcal{I}_{p+1}^{pyr(k)}$ by collecting images from sequence $\{\mathcal{I}_q^F : 1 \leq q \leq n\}$
  **for** $k = N_{pyr}-1$ *to* 0 **do**
    $\mathcal{I}_1 = \mathcal{I}_p^{pyr(k)}$, $\mathcal{I}_2 = \mathcal{I}_{p+1}^{pyr(k)}$;
    **if** $k == N_{pyr}-1$ **then**
      Set $\boldsymbol{w}_p = \mathbf{0}$;
    **else**
      Upsample $u$, $v$ to the current level;
    **if** $k \neq N_{pyr}-1$ **then**
      Update $\mathcal{I}_1 = warping(\mathcal{I}_1, \boldsymbol{w}_p)$;
    Initialize $\Delta\boldsymbol{w}_p$ and $\Delta\hat{\boldsymbol{w}}_p$ to zero with $size == size(\mathcal{I}_1)$;
    **for** $w = 1$ *to* $N_{warp}$ **do**
      **for** $i = 1$ *to* $N_{it}$ **do**
        **if** *i==1* **then**
          Initialize $\Lambda_{\mathcal{W}}$ for $\mathcal{W} = u$ or $v$;
        Update $\hat{u}$ and $\hat{v}$;
        Evaluate $\hat{u}^+$, $\hat{u}^-$, $\hat{v}^+$, $\hat{v}^-$ using $\hat{u}$, $\hat{v}$ and $\Lambda_{\mathcal{W}}$;
        Update $\Lambda_{\mathcal{W}}$ for $\mathcal{W} = u$ or $v$; Estimate $\Delta\boldsymbol{w}_p$ using updated $\Lambda_{\mathcal{W}}$;
      Update $\boldsymbol{w}_p = \boldsymbol{w}_p + \Delta\boldsymbol{w}_p$;
      Update $\mathcal{I}_1 = warping(\mathcal{I}_1, \Delta\boldsymbol{w}_p)$;
      Initialize $\Delta\boldsymbol{w}_p$ to zero;
**for** $z = 1$ *to* $n-1$ **do**
  Calculate $\mathcal{V}_z$, the $4D$ vector using $\mathcal{I}_z^{mask}$ and $\boldsymbol{w}_z$;
$\mathcal{I}_{ref} = I_1$;
$Mixed\_data = \{\mathcal{I}_{ref}, \{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_{n-1}\}\}$;

the proposed algorithm with the existing techniques, their fire detection results have been cited from the respective publications. The existing models are chosen for comparison purpose based on the types of dataset and features such as image data [22]–[24], GIS (Geographic Information System) data [25], temperature, $CO_2$ and relative humidity [26] etc. Table I illustrates that the propounded algorithm outperforms these existing techniques.


Fig. 3. Sample reference images used in fire detection: First row contains fire images, and second row contains non-fire images.

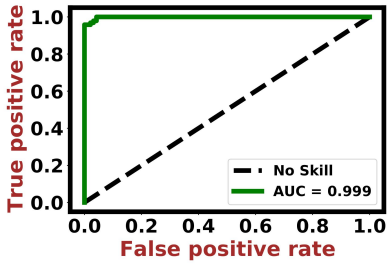


Fig. 4. ROC curve corresponding to the proposed framework.

Last experiment illustrates the significance of different components of the proposed algorithm through an ablation study. In order to perform this ablation study, the algorithm is considered in three different forms. In first form, the fire detection is performed by using the color maps only. A color map is an RGB image representation of an optical flow field. In a color map, different colors correspond to distinct flow directions and color intensities describe the flow magnitudes [15]. The second form performs fire detection by using the fire color enhanced images only. In third form, the training and testing is conducted only with raw image frames. The results are described in Table II. Thus, the novelty of model is justified.

## IV. Conclusions and Future Work

The proposed work introduced a fire detection framework based on the fusion of static and dynamic features of fire. The static features were considered as color, texture and shape of fire, whereas the dynamic features were estimated using the motion edge preserved LSS based fractional order fire optical flow. Optical flow fields are

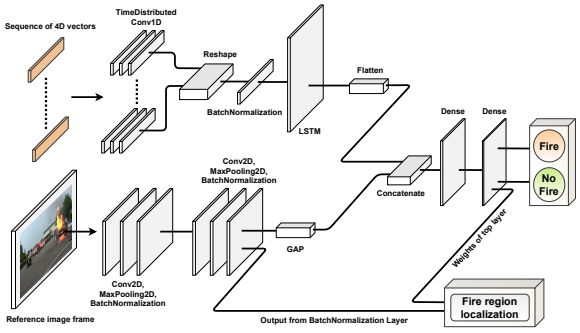


Fig. 5. Architecture of the proposed mixed data CNN-LSTM model.

TABLE I
COMPARISON BETWEEN THE PRESENTED ALGORITHM AGAINST THE STATE-OF-THE-ART ALGORITHMS.

| **Algorithm** | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| **Proposed** | **0.97** | **0.93** | **1.00** | **0.96** |
| DBN [22] | 0.94 | 0.93 | 0.96 | 0.94 |
| DLBN [23] | 0.89 | 0.98 | 0.81 | 0.89 |
| ResNet50 model [24] | 0.89 | 0.81 | 1.00 | 0.89 |
| ResNet101 model [24] | 0.95 | 0.95 | 0.93 | 0.94 |
| Catry et al. [25] | 0.80 | – | – | – |
| Yan et al. [26] | 0.83 | – | – | – |
| Liu et al. [27] | 0.40 | – | – | – |

TABLE II
ABLATION STUDY OF THE PROPOSED ALGORITHM.

| | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| **Proposed algorithm** | **0.97** | **0.93** | **1.00** | **0.96** |
| Color maps | 0.72 | 0.83 | 0.67 | 0.74 |
| Fire enhanced images | 0.60 | 0.33 | 0.33 | 0.33 |
| Raw images | 0.50 | 0.33 | 0.67 | 0.40 |

reduced into $4D$ feature vectors which describe the fire motion related information. This flow field to $4D$ vector transformation provided data dimensionality reduction that helps in mitigating overfitting. Further, a mixed dataset has been prepared, which contains reference image frames and their corresponding sequences of $4D$ feature vectors. The classification accuracy obtained for the given algorithm is 0.97, which is quite satisfactory. The presence of fire is also illustrated through probability scores. ROC curve demonstrated a high AUC score that indicated a good classification performance corresponding to different test samples. A detailed comparison established the validity of the proposed algorithm. Finally, an ablation study emphasized the importance of the proposed framework. The processing time taken by the CNN-LSTM model is 124 ms/input sample, while that for deep learning with raw data is 92 ms/input sample. Yet, the ablation study demonstrated that the proposed model provides twice more accurate results. Thus, in future, this algorithm can be further improved by using four color theorem to estimate more accurate fire flow fields.

## ACKNOWLEDGMENT

Pushpendra Kumar acknowledges the support of NBHM, Mumbai for grant no. 02011/24/2021 NBHM (R.P)/ R&D-II/8669. The author Muzammil Khan acknowledges thankfully to MHRD, New Delhi, Government of India. Also, the author Pushpendra Kumar acknowledges the support of SERB, New Delhi for grant no: EEQ/2020/000154.

## REFERENCES

[1] D. P. Nolan, *Handbook of fire and explosion protection engineering principles: for oil, gas, chemical and related facilities*. William Andrew, 2014.

[2] N. F. P. Association, "Smoke alarm (smoke detector) troubleshooting," *https://www.nfpa.org*, 2022.

[3] S. Gupta, M. Khan, and P. Kumar, "Prediction of fire signatures based on fractional order optical flow and convolution neural network," in *International Conference on Computer Vision and Image Processing*, 2023, pp. 308–321.

[4] R. Xu, H. Lin, K. Lu, L. Cao, and Y. Liu, "A forest fire detection system based on ensemble learning," *Forests*, vol. 12, no. 2, p. 217, 2021.

[5] S. Li, Q. Yan, and P. Liu, "An efficient fire detection method based on multiscale feature extraction, implicit deep supervision and channel attention mechanism," *IEEE Transactions on Image Processing*, vol. 29, pp. 8467–8475, 2020.

[6] Q. An, X. Chen, J. Zhang, R. Shi, Y. Yang, and W. Huang, "A robust fire detection model via convolution neural networks for intelligent robot vision sensing," *Sensors*, vol. 22, no. 8, p. 2929, 2022.

[7] M. Mueller, P. Karasev, I. Kolesov, and A. Tannenbaum, "Optical flow estimation for flame detection in videos," *IEEE Transactions on image processing*, vol. 22, no. 7, pp. 2786–2797, 2013.

[8] B. K. Horn and B. G. Schunck, "Determining optical flow," *Artificial intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.

[9] M. Zhai, X. Xiang, N. Lv, and X. Kong, "Optical flow and scene flow estimation: A survey," *Pattern Recognition*, vol. 114, p. 107861, 2021.

[10] D. Chen, H. Sheng, Y. Chen, and D. Xue, "Fractional-order variational optical flow model for motion estimation," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1990, p. 20120148, 2013.

[11] P. Kumar, S. Kumar, and B. Raman, "A fractional order variational model for the robust estimation of optical flow from image sequences," *Optik*, vol. 127, no. 20, pp. 8710–8727, 2016.

[12] P. Kumar, S. Kumar, and R. Balasubramanian, "A fractional order total variation model for the estimation of optical flow," in *Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, 2015, pp. 1–4.

[13] M. Khan and P. Kumar, "A nonlinear modeling of fractional order based variational model in optical flow estimation," *Optik*, vol. 261, p. 169136, 2022.

[14] S. Arora, T. Mathur, S. Agarwal, K. Tiwari, and P. Gupta, "Applications of fractional calculus in computer vision: a survey," *Neurocomputing*, vol. 489, pp. 407–428, 2022.

[15] M. Khan and P. Kumar, "A level set based fractional order variational model for motion estimation in application oriented spectrum," *Expert Systems with Applications*, p. 119628, 2023.

[16] M. Rahimzad, A. Moghaddam Nia, H. Zolfonoon, J. Soltani, A. Danandeh Mehr, and H.-H. Kwon, "Performance comparison of an lstm-based deep learning model versus conventional machine learning algorithms for streamflow forecasting," *Water Resources Management*, vol. 35, no. 12, pp. 4167–4187, 2021.

[17] "Fire segmentation image dataset," in *Automatically Generated Fire Segmentation Image Dataset*, https://www.kaggle.com/datasets/diversisai/fire-segmentation-image-dataset.

[18] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Two deterministic half-quadratic regularization algorithms for computed imaging," in *International Conference on Image Processing*, vol. 2, 1994, pp. 168–172.

[19] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International journal of computer vision*, vol. 50, pp. 271–293, 2002.

[20] A. Marchaud, "Sur les d&iv&es et sur les differences des fonctions de variables rielles," *J. Math. Pures Appl*, vol. 6, p. 337425, 1927.

[21] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical imaging and vision*, vol. 20, pp. 89–97, 2004.

[22] A. S. Pundir and B. Raman, "Deep belief network for smoke detection," *Fire technology*, vol. 53, pp. 1943–1960, 2017.

[23] Y. Wu, M. Chen, Y. Wo, and G. Han, "Video smoke detection base on dense optical flow and convolutional neural network," *Multimedia Tools and Applications*, vol. 80, pp. 35 887–35 901, 2021.

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Computer vision and pattern recognition*, 2016, pp. 770–778.

[25] F. X. Catry, F. C. Rego, F. L. Bação, and F. Moreira, "Modeling and mapping wildfire ignition risk in portugal," *International Journal of Wildland Fire*, vol. 18, no. 8, pp. 921–931, 2009.

[26] X. Yan, H. Cheng, Y. Zhao, W. Yu, H. Huang, and X. Zheng, "Real-time identification of smoldering and flaming combustion phases in forest using a wireless sensor network-based multi-sensor system and artificial neural network," *Sensors*, vol. 16, no. 8, p. 1228, 2016.

[27] Y. Liu, Y. Gu, G. Chen, Y. Ji, and J. Li, "A novel accurate forest fire detection system using wireless sensor networks," in *International Conference on Mobile Ad-hoc and Sensor Networks*, 2011, pp. 52–59.