



DroneRTEF: development of a novel adaptive framework for railroad track extraction in drone images

Aradhya Saini¹ · Dharmendra Singh^{1,2}

Received: 21 May 2020 / Accepted: 18 May 2021 / Published online: 29 June 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

Railroad track health monitoring is a challenging yet important task as it affects the safety of railroad systems. Railroad track extraction presents an immediate advantage during railroad inspections in an efficient and cost-effective manner. At present, human inspectors, inspection trains, and rail-mounted vehicles equipped with cameras are prevalent image acquisition systems (IAS) in the track extraction module. However, these IAS face various challenges such as high operability cost, railroad closed for normal traffic, and inaccessibility to certain geographical locations. In such scenarios, drones act as effective IAS. Therefore, this paper presents a novel and adaptive railroad track extraction framework for drone images (DI) captured under uneven illumination, at different drone flight heights, with varying rail line orientations, and in complex railroad environments. We termed this framework as DroneRTEF. This work primarily focuses on two aspects of drone-based railroad track images: image enhancement and image analysis. With regard to the first aspect, a global image enhancement algorithm named adaptive colour space-based masking (ACSM) is developed to enhance railroad track images and identify rail lines. The rail lines and background can be highlighted and homogenized, respectively, in DI captured under various sunlight intensity using ACSM due to its illuminance independence. With regard to the second aspect, the Hough parameter space analysis-based novel Hough transform-ground sample distance (HT-GSD) method is presented in this paper. The proposed HT-GSD method emphasizes on rail line detections at varying line orientations and different flight heights. The track extraction is then performed by a coordinate transformation technique. The approach has been successfully tested and validated on various DI. The efficacy of our framework for rail line detection is identified by comparing it with other line detection model. Performances of these methods are tested using metrics such as precision, recall and accuracies of the detections. Results obtained show that our method is superior to another model. Therefore, DroneRTEF is an efficient and feasible method for railroad track extraction in DI.

Keywords Railroad · Drone Images · Image Processing · Adaptive Thresholding · Ground Sample Distance

1 Introduction

Railroad tracks are a prime focus in the railroad environment as they largely benefit the community through the provision of passenger comfort, effective railroad operations, and

train speed modulations. A railroad track comprises of two rail lines, the track space between these lines and the lateral area (right and left side of the rail lines) comprising of the ballast as shown in Fig. 1a. Railroad track monitoring is essential to ensure the effective maintenance of operational rail lines in the safety-critical railroad infrastructure. Railroad track extraction is an important pre-requisite step in the track monitoring module. Track extraction is a crucial step as it helps to rapidly segment the railroad track (target area). This extraction of track area facilitates tasks such as inspection and detection of rail surface defects [1, 7], automated fastener classification and defect detection [4], train driver assistance and obstacle identification [2, 12] provision of autonomous train control systems [3], computation

✉ Dharmendra Singh
dharmfec@gmail.com

Aradhya Saini
aradhya.saini91@gmail.com

¹ Department of Electronics and Communication Engineering,
Indian Institute of Technology Roorkee, Roorkee,
Uttarakhand, India

² RailTel - IIT Roorkee Center of Excellence
in Telecommunication, Roorkee, Uttarakhand, India

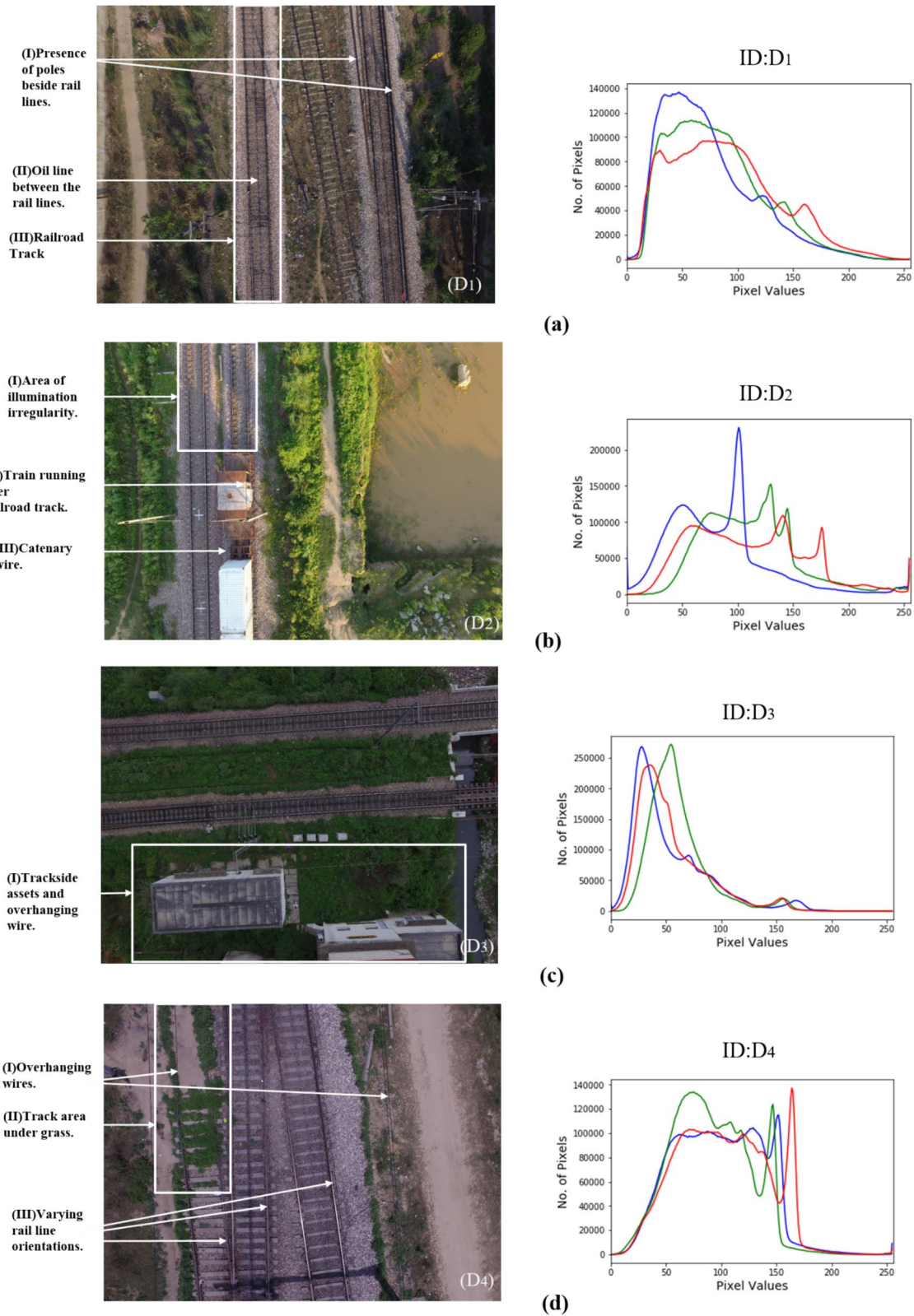


Fig. 1 a Study Image 1 b Study Image 2 c Study Image 3 d Study Image 4 with their corresponding image histogram in RGB

of asset-sighting distance [13] and vegetation condition monitoring [14] in a convenient, and cost-effective manner.

There has been an advent of various railroad track IAS in conjunction with research on models using image processing and computer vision techniques for track detection and extraction. Dynamic Programming is used for the extraction of railroad track space through video sequences captured by a camera placed in front of a train engine [2]. A combination of image processing algorithms and line segment detector (LSD) is used for track detection in images captured through cameras hung below a train coach [4]. The camera is fixed behind the windshield and the track extraction is carried out using techniques such as priori shape model and gradient information [5, 6]. The semantic segmentation method is computed in the visual-based track inspection (VITS) collected images in order to extract rail tracks and locate ROIs using U-Net [7]. In fact, there are several other methods which include HOG feature extraction for track detection in images captured through a camera placed in the vehicle [8], histogram-based track extraction (HBTE) where IAS is under a test train [9], geometry constraints-based track extraction in images captured from the camera installed in the front of a train [10], track extraction based on projection profile (TEBP) in rail track images captured by a camera fixed on a train [11]. Also, the target tracks have been segmented by feature analysis in the hue channel [1]. However, this method uses a portable track defect vision inspection prototype for image acquisition. Angle alignment measure is computed for real-time railroad extraction in images captured using a mobile phone [12]. Rail track detection is performed in image data captured from a rail-mounted vehicle camera using cubic Bezier curves [13]. HOG and mean-shift clustering are used for the detection of rails in images captured from DSLR camera placed within a trolley [14].

However, these aforementioned IAS have varying inevitable drawbacks which include high cost, limited detection range, railroad becomes closed to normal train traffic during operations, and inaccessibility to remote geographical locations. Therefore, researchers are exploring drone-based image acquisition that may come to rescue during such scenarios. Usage of drones or unmanned vehicles for railroad infrastructure monitoring is the latest trend and currently, the camera sensors are the richest data sources [15–17]. Drones are lightweight UAVs (Unmanned Aerial Vehicles) which provide various advantages such as cost-effectiveness, efficient image acquisition while capturing track images without blocking the railroad traffic, ease of control, and flexibility while aiming areas inaccessible to human inspectors, inspection trains or rail-mounted vehicles.

Although the drone-based image acquisition module offers various advantages, still research on drone-based railroad track extraction faces the following challenges:

1. Due to different sunlight conditions (sunny, partially sunny/cloudy), there might be illumination inconsistency. Also, partial occlusion due to infrastructures (such as overhead catenaries along with respective support structures), rail reflectance properties, shaking of the drone along with other environmental aspects leads to uneven and low contrast and brightness of the captured DI. This necessitates an adaptive method for track extraction in uneven illumination.
2. Rail line positions vary in drone images. The angle of the HD camera installed on the drone is sensitive to various environmental factors (which includes turbulence and wind) as well as operators. Although drone can balance itself through GPS flight mode still positions of rail lines are extremely variable in drone images. Hence, these variances in the rail line positions lead to complex track extraction scenarios.
3. Flying drones at various flight heights may lead to capturing different views of the same area/different areas. Also, the presence of poles lying beside rail lines, oil lines between rail lines, track area under grass, train running over railroad tracks are some complex railroad environments in the captured DI which make track extraction a difficult task.

Due to these challenges, it is difficult to exercise the above discussed various IAS-based track detection and extraction models on drone railroad track images. This acts as the motivation to overcome these shortcomings and propose a novel railroad track extraction framework that can locate the rail lines and extract the railroad track from drone images acquired in uneven sunlight intensity, at varying flight heights, and with different rail line positions.

1.1 Contributions

The main contributions of the proposed work are summarized as follows:

1. We propose a novel and adaptive railroad track extraction method for drone images termed as DroneRTEF. This framework is based on computing colour feature-based adaptive image enhancement of DI to locate rail lines and Hough parameter space-based novel image analysis method for track extraction. To the best of our knowledge, this is the first attempt to examine the hybrid approach of the colour feature and Hough parameter space analysis for railroad track extraction in drone images captured at varying flight heights, with different rail lines orientations, and in uneven illumination.
2. Due to the unavailability of a standard railroad track dataset comprising of drone images captured in diverse

railroad environments, the image acquisition is also performed to capture datasets of track DI for the research work.

3. The performance of the proposed rail line detection module is evaluated and validated with different experiments upon 3 different metrics and is compared with another rail line detection algorithm.

The organization of the article is as follows: In Sect. 2, algorithms related to railroad track detection and extraction are discussed. Image acquisition and dataset description are presented in Sect. 3. The theoretical background is presented in Sect. 4. The proposed novel adaptive railroad track extraction framework is discussed in Sect. 5. In Sect. 6, we present experimental results and discussions. In this section, we have also discussed the evaluation and validation of the proposed method. Finally, this paper is concluded along with future work in Sect. 7.

2 Related work

A typical approach for railroad track extraction can be carried out using image processing algorithms [18]. These algorithms can be generalized into two steps: image enhancement and image analysis. The image enhancement method is used to pre-process the track drone image and then segment it in order to highlight the rail line feature information. This is advantageous as it helps to eliminate or weaken the interference of background information and identify rail lines. The image analysis step is performed by extracting the features (shape features/geometric features) of the identified rail lines. Several studies have been previously reported in the literature on railroad track image enhancement and analysis for track detection and subsequent extraction in DI.

The local weber-like contrast(LWLC) algorithm has been used for the enhancement of rail images [19]. However, this method is performed on a local window after rail track extraction step. Therefore, it is unable to give adequate insight into if LWLC is an effective global image enhancement step for track extraction in high-altitude drone images. Colour pattern template matching is computed for constraining rail search locations by detecting railroad profile using ROI technique [20]. Although it has been noted that these results are only possible if the template of railroad of interest is used. Histogram equalization (HE) is also a popular method for image enhancement in global lighting conditions [19, 20, 21]. However, it poses disadvantages such as loss of detail information from the image and noise amplification [19]. Mask bank technique is found advantageous for locating railroad tracks in aerial images [22]. However, this mask bank contains masks that run only from top to bottom. Also, this method works in two phases: training and

detection where further testing is required to prove functionality. Nonetheless, masks prove to be versatile and provide a quick way to access information. Also, colour is a distinctive and powerful feature for identification in the railroad environment [1, 23, 24]. It has been observed that colour spaces in conjunction with masks are quite helpful for rail line identification in DI while working in complex railroad environments [22, 25]. Through this critical analysis, it is observed that colour space-based masking seems an effective image enhancement method for the segmentation and identification of rail lines in railroad track DI captured in uneven illumination.

In earlier studies, a range of image processing techniques has been developed for image analysis of rail lines in DI. Linear regression model is used for detecting valid rail lines [26]. However, the regression method is less robust to noisy data [27]. Several other rail line detection algorithms have also been discussed such as correlation matching, Touzi filters, CNN-based semantic segmentation [28]. All these methods hint at using a prior infrastructure ground coverage mask for improved performance [28]. Canny detector and k-NN Mean Euclidean distances classifier have also been used for the detection of rail tracks [29]. In this work, k-NN classifier faces challenges as it is slow, consumes a lot of memory, and requires rail edges for training. Deep drone vision uses 2-step CNN for computing rail track detection [30]. However, this method is computationally intensive as it requires training and the dataset faces limitations due to its inability to accommodate varying degrees of illuminations at different times of the day. Hough transform (HT) is based on the duality of line and can be used for railroad track extraction [19, 31]. Hough transform and cumulative grey value of each pixel column have been implemented for track extraction in drone images captured only at 30 m [19]. Several variants of this transform have already been discussed in the literature [17, 31, 32]. Also, Probabilistic Hough transform (PHT) [31] is one of the variants. However, PHT cannot be used instead of HT pertaining to the computational complexity of PHT [33]. Hough transform (an under constraint line fitting algorithm) serves advantageous over other line detection algorithms such as RANSAC (constraint bound) [6, 17] and Least Squares Fit (over constraint). Also, the Hough transform model can handle a higher percentage of outlier points and is best suited for highly noisy data [34]. Therefore, inspired by these successes, HT is a suitable solution for geometrical representation of Hough parameter space of rail lines oriented at varying angles. Nonetheless, the detection of rail lines in drone images captured at different altitudes is another essential step for track extraction at different flight heights. In order to fulfil this aim, ground sample distance (GSD) method [35] is useful as changing the altitude of the drone alters the GSD of the image which in turn helps to detect valid rail lines (rail line pairs) at different

flight heights and subsequently extract the railroad track. Alongside, the other advantages of the method include: global image transformation of drone images which provides adequate insight into track extraction, template of the railroad of interest is not required, ground truth is not required at any stage of the framework, training is not required which makes the algorithm less computationally intensive.

Therefore, our aim is railroad track extraction in DI captured with uneven illumination, varying rail line orientations, and at different flight heights. To achieve this aim, we propose the novel DroneRTEF framework in this paper as follows:

1. Present a global image enhancement algorithm termed as adaptive colour space-based masking (ACSM). This algorithm is adaptive as it computes colour space to enhance DI captured under uneven natural light and then creates track mask to segment these images and identify rail lines in them.
2. Propose Hough parameter space analysis-based novel Hough transform-ground sample distance(HT-GSD) algorithm for rail line detection and track extraction. In this image analysis method, we first detect valid identified rail lines at varying orientations and flight heights, which then augments segmenting the required railroad track area.

3 Image acquisition and dataset description

3.1 Study area and data sets used

In this work, the image acquisition has been carried out using DJI phantom quadcopter. This drone constitutes a high definition 4 K resolution RGB colour camera, with an integrated GPS unit in order to capture geotagged standardized RGB (sRGB) images each of size 3000×4000 ($I_h \times I_w$) pixels. The acquired track datasets have about 415 track DI, captured during the course of various flights taken over railroad tracks at different track locations near Roorkee,

Haridwar district, India. These DI are field acquired frame by frame at varying flight heights and with different rail line orientations, different locations, different time durations and have managed to capture gamut of variations in complex railroad environments under different sunlight illumination.

Some of these railroad track images have been selected as study images as shown in Fig. 1. These images are taken from the field captured track datasets for evaluating the proposed framework. The details of these study images, as mentioned in Table 1, include their data ID, date of acquisition (DOA), flight height (F_h) (in metres (m)), central latitude and longitude, GSD, illumination conditions and the total number of images in the dataset from which the corresponding study image has been selected. The railroad environment scenarios in which these study images are captured are described as follows:

Scenario 1: Study image 1(D_1) as shown in Fig. 1a is captured at 4:48p.m. at a flight height of 22 m on a sunny day with little occlusion and constitutes poles beside tracks and oil lines.

Scenario 2: Study image 2 (D_2) shown in Fig. 1b is captured on a bright sunny day at 6:04 p.m. at 25 m flight height and showcases illumination irregularity with partial occlusion on the track. The other components constitute train running over a railroad track, catenary wires, dense bushes along the track.

Scenario 3: Study image 3(D_3) in Fig. 1c is also captured at 2:50 p.m. at 25 m. Even though images both in Figs. 1b and c are captured at the same height still Fig. 1c image showcases a low brightness (dark) environment. It also comprises of assets along the track and overhanging catenary wires.

Scenario 4: Study image 4(D_4) in Fig. 1d is captured at 11:53 a.m. at 11 m on a little brighter day and showcases varying rail line orientations along with track areas covered with vegetation and shadowed by catenary wires.

Table 1 Dataset description

SN	ID	F_h (m)	DOA	Central Lat/Long	GSD (cm)	Illumination Scenario	No. of images in dataset
(1)	D_1	22 m	10–05-2018	29°51'0.2884"N/ 77°52'52.7952"E	0.89	Sunny, Little occlusion	56
(2)	D_2	25 m	30–08-2017	29°49'38.5572"N/ 77°55'31.3284"E	1.06	Quite Sunny, Partially occluded	201
(3)	D_3	25 m	07–07-2017	29°51'13.4073"N/ 77°52'11.9761"E	1.06	Dark	63
(4)	D_4	11 m	25–04-2017	29°46'3.3492"N/ 78°0'35.0172E	0.47	Little bright	96

3.2 Challenges faced while processing DI

As discussed in [Introduction \(Sect. 1\)](#), DI show variations in terms of railroad tracks captured under different sunlight intensity, with varying rail line orientations, at different flight heights and in complex railroad environment scenarios. These variations make it challenging to process DI for rail line segmentation and identification. In such scenarios, the following observations have been made:

1. Image characteristics vary under challenging railroad environment imaging conditions. Each of the study image(D_1, D_2, D_3, D_4) plots in Figs. 1a–d showcase the pixel intensity count for red, green, and blue channels of the corresponding DI in the form of red, green, and blue histograms, respectively. It can be observed that plots of different DI exhibit different histogram behaviours when viewed in RGB colour space.
2. Rail lines are a constant feature in all the DI. However, it has been inferred from the histogram analysis that no uniform histogram pattern indicating the presence of rail lines has been observed among the colour channel(s) of different images. This is because the rail line area is very small and the external effects (such as illumination) are quite inconsistent. Therefore, it is difficult to accurately segment and identify the rail lines from the background using a histogram-based thresholding method that may be useful for the creation of a railroad track mask comprising of only the rail lines.
3. The aim is to segment and identify only rail lines in DI. For this purpose, various thresholding algorithms like Otsu's method[36], entropy-based thresholding [37] have also been attempted. However, it has been observed that it is difficult to apply a common thresholding algorithm for all images in order to obtain a track mask.

Therefore, this makes it a challenging task to devise an algorithm to create a track mask in order to segment and identify the rail lines in DI captured under varied image acquisition scenarios in the railroad environments.

4 Theoretical background

In view of the discussion in [Sect. 2](#), it is essential to address non-uniform illumination, rail line orientations, and change in flight heights in order to detect rail lines and extract railroad tracks from DI. As discussed in [Sect. 3](#), there is no uniform histogram pattern in order to threshold only the rail lines(rail line area) in images and even other thresholding algorithms do not hold good results. Therefore, colour space transformations are essential for track mask creation to identify only the rail lines in DI with uneven illumination. Also,

the analysis of the GSD is essential as it facilitates track extraction at various flight heights and rail line orientations. To fulfil the aforementioned aims, we have critically analysed the following parameters.

4.1 Colour space transformations

Colour is a fundamental descriptive property of objects in various object identification and segmentation frameworks [38–40]. However, perception subjectivity and extensive need of the application draw an uncertainty over which colour space fits a certain application during the segmentation process. In our work, no prior preference has been given to any specific colour models (colour spaces). The human eye is considered more sensitive to the colour changes than the grey-scale variations. Hence, here, we have concentrated on the following colour models: RGB, HSV, $L^*a^*b^*$, YCbCr [40]. The colour space transformation functions are described in [Table 2](#).

RGB colour model is widely used for the storage of digital data in railways [44]. The colour space is additive in nature and is defined by the sum combination of three primary colours which are represented by red (R), green (G), blue (B) chromaticities. The aim is to obtain a railroad track mask which comprises of only the rail lines while the rest of the background (comprising of the vegetation, sleepers, ballast, poles lying beside the rail lines) in DI is removed. However, the RGB colour space holds disadvantages (1) It is device-dependent (2) It is associated with an inherent problem of mixing the luminance and the chrominance data(present in all three channels) [41, 42].

Hue, saturation and value (HSV) are a cylindrical transformation [23] where H,S,V components denote dominant wavelength, purity of the colour and intensity value, respectively. The H and S are invariant to shading and shadows [40]. Also, hue(H) component is found to be independent of illumination conditions[1] and specularities [40]. The equations for RGB to HSV space transformation function T are stated in [Table 2](#). Even though the HSV colour space is device-dependent but it presents with an advantage that the colour information is intact only in one channel (H). It also presents the advantage of the separation of luminance from chrominance information[42]. Also, the colour space provides an advantage of easily selecting the desired hue which can be slightly modified by adjusting the saturation and value [42, 45]. HSV histograms for images are as shown in [Appendix A](#).

Luminance-based $L^*a^*b^*$ (CIELAB or Lab) colour space stems from the CIE colour model and is considered as perceptually uniform colour space [24, 40]. The $L^*a^*b^*$ colour space is device independent, and the colour information is encoded in only two colour channels (a^* and b^*). The RGB

Table 2 Colour space transformations

Colour space	Conversion formulae
Standardized or Normalized RGB(rgb/sRGB) [41]	$r = \frac{R}{R + G + B}$ $g = \frac{G}{R + G + B}$ $b = \frac{B}{R + G + B}$ $r + g + b = 1$
HSV [42]	$V \leftarrow \max(R, G, B)$ $S \leftarrow \begin{cases} \frac{V - \min(R, G, B)}{V} & \{if V \neq 0\} \\ 0 & \{otherwise\} \end{cases}$ $H \leftarrow \begin{cases} 60(G - B)/(V - \min(R, G, B)) & \{if V = R\} \\ 120 + 60(B - R)/(V - \min(R, G, B)) & \{if V = G\} \\ 240 + 60(R - G)/(V - \min(R, G, B)) & \{if V = B\} \end{cases}$
CIE $L^*a^*b^*$ [42]	$L^* = \begin{cases} 116(\frac{Y}{Y_n})^{1/3} - 16 & \{if \frac{Y}{Y_n} > 0.008856\} \\ 903.3(\frac{Y}{Y_n}) - 16 & \{if \frac{Y}{Y_n} \leq 0.008856\} \end{cases}$ $a^* = 500 * (f(\frac{X}{X_n}) - f(\frac{Y}{Y_n}))$ $b^* = 200 * (f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n}))$
YCbCr [41, 43]	$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.24 & 0.50 & 0.098 \\ -0.14 & -0.291 & 0.439 \\ 0.43 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$

to $L^*a^*b^*$ space conversion is carried out using the respective transformation function T stated in Table 2.

YCbCr colour space is considered perceptually uniform with an approximation [40]. The YCbCr colour space is device-dependent and the colour information is encoded in both Cb and Cr (only two colour channels). Additionally, it is advantageous as it separates the RGB into chrominance and luminance information[42]. The RGB to YCbCr conversion formulae T is stated in Table 2.

4.2 Edge characterization and geometrical representation

The prerequisite to rail line detection process is creation of an edge map (discussed in Sect. 5.2). This edge map comprises the detected edges from the railroad track mask by using an edge detection algorithm. Such an edge detection algorithm should be able to provide localization and detect right number of rail line edge responses. Localization corresponds to detection of rail line edges at the right location.

The accurate shape and geometrical representation of edge map in Hough parameter space are necessary for rail line detection as many objects in the railroad environment

are characterized by straight lines. It is essential to determine which edge belongs to which line and record all possible lines upon which edge points (of the edge map) lie (discussed in Sect. 5.3).

4.3 GSD calculation

GSD is defined as the distance between the pixel centres of the camera sensor as measured on the ground and indicates how big each pixel is on the ground. As the flight height changes, the corresponding pixel size changes (GSD in cm), as shown in Table 1. This change in the pixel size leads to a subsequent increase or decrease in the number of pixels between two valid rail lines of the rail line pair. Therefore, the calculated number of pixels is essential during rail line pair selection (discussed in Sect. 5.3.3). Consequently, GSD helps in railroad track extraction from DI at any given flight height F_h and makes the algorithm adaptive. The calculated distance, GSD , is a function of the flight height, sensor dimensions, and image measurements (in pixels). For any pixel, the GSD mapping on the ground is represented by the pair GSD length (GSD_l) and GSD width (GSD_w), which is computed as in (1), (2) and (3) [35]:

$$GSD_h = \frac{F_h \times S_h}{Flen \times I_h} \tag{1}$$

$$GSD_w = \frac{F_h \times S_w}{Flen \times I_w} \tag{2}$$

$$GSD = GSD_h \times GSD_w \tag{3}$$

Here, F_h denotes the flight height above ground which is evaluated from Table 1, $Flen$ represents the focal length of the camera sensor, S_w and S_h are the sensor width and height, respectively, I_h and I_w are the image height and width, respectively (in pixels).

5 The proposed algorithm

The primary objective of the proposed algorithm DroneRTEF is to develop a framework for railroad track extraction from DI captured under uneven sunlight intensity, with different rail line orientations and at varying flight heights in the railroad environment. This aim is facilitated by the capability of the algorithm to characterize the colour, shape, and geometric features of rail lines for railroad track extraction. The proposed DroneRTEF framework is divided into two stages:

1. ACSM algorithm is developed to identify rail lines in DI captured in uneven illumination. In order to achieve this aim, an adequate colour space is selected and then thresholding is performed to create a railroad track mask. This mask segments DI to identify rail lines in them.
2. HT-GSD algorithm is a novel method for Hough parameter space-based characterization of shape and geometric features of rail lines. This method uses Hough param-

eters ρ and θ , to detect the identified rail lines and then perform rail line pair selection. Coordinate transformation is then computed for extraction of the railroad track from DI captured at different orientations and flight heights.

The complete flow diagram of the proposed algorithm is as shown in Fig. 2, and the steps are discussed in Sect. 5.1, 5.2 and 5.3.

5.1 Adaptive Colour Space-based masking (ACSM)

In order to create a railroad track mask to identify rail lines in DI, we propose a global image enhancement algorithm adaptive colour space-based masking (ACSM) as described in Algorithm 1 in Fig. 3. As mentioned in Sect. 4.1, four colour spaces: RGB, HSV, $L^*a^*b^*$, YCbCr have been proposed for railroad track mask creation in DI. The visual analysis of track masks is discussed in Sect. 5.1.1.

The track mask creation method is devised to perform rail line identification during uneven sunlight illumination in DI. This work is challenging as:

1. Rail lines occupy small areas and are difficult to locate uniformly in DI.
2. A greater variation in the range of the grey values has been observed in the global scope. This may be due to the reflectance properties of the railroad surfaces and uneven illumination.

The ACSM method is devised to overcome these challenges and create a track mask. Each step is given in Algorithm 1 and is performed as follows:

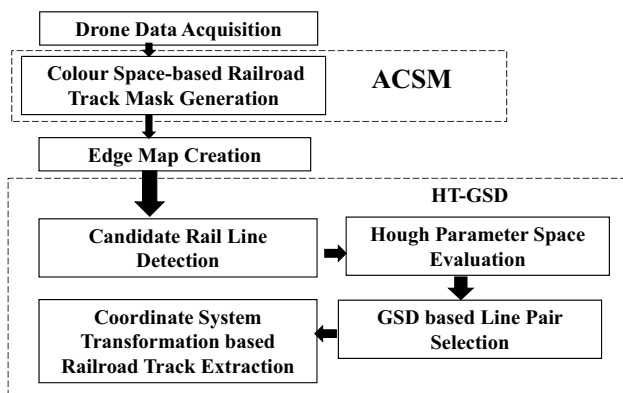


Fig. 2 DroneRTEF framework

Algorithm 1: Adaptive Colour Space Based Masking(ACSM)

Input: Drone Image D_i
Output: Railroad Track Mask M_i

1. Obtain O_i by computing T over D_i ;
2. for $j:=1:n$:
3. if $low_limit_{color_space}(j) < o_j < upper_limit_{color_space}(j)$.
4. return $O_{j|thresh} := O_j$
5. $O_{i|thresh} := O_{1|thresh}, O_{2|thresh}, \dots, O_{n|thresh}$
6. $M_i := O_{i|thresh}$

Fig. 3 ACSM algorithm

1. *Global Image transformation:* In this method, we perform evaluations of different colour spaces based on their discriminative power and their ability to provide photometric invariance in DI. This technique applies different colour space transformations T (in Sect. 4.1) using Eqs. 4,5 and 6, with the aim to enhance DI D_i .

$$o_j = t_j(l_1, l_2, l_3, \dots, l_n) \quad j = 1, 2, 3, \dots, n \tag{4}$$

$$T = \{t_1, t_2, t_3 \dots t_n\} \tag{5}$$

$$O_i = T[D_i] \tag{6}$$

where O_i is transformed(enhanced) colour output drone image, D_i is colour input drone image, o_j and l_j denote colour components of the transformed colour output image $O_i(x, y)$ and colour input image $D_i(x, y)$ at any point (x, y) , t_j implies colour mapping function, n equals number of colour components and n functions, t_j , combine to implement one transformation function, T .

2. *Track Mask creation:* In this method, the transformed image O_i is now segmented to identify rail lines using Eq. (7). The colour threshold-based segmentation is performed in each of the colour spaces on corresponding o_j components in order to obtain mask M_i . The thresholds used in equation (7) are determined through critical analysis of different values for the lower and upper limit in the range of intensity values of colour channels. Thus, in this mask only the rail lines may be visible and the rest of the background in DI(such as catenaries, assets

along the track, track signs, vegetation along-track, ballast, sleepers, pole lines) may be eliminated.

$$M_i = O_{i|thresh} \tag{7}$$

where M_i denotes track mask, $O_{i|thresh}$ denotes thresholded O_i image (o_j components) in the range $\in [\text{low_limit}_{color_space}, \text{upper_limit}_{color_space}]$, $[\text{low_limit}_{color_space}, \text{upper_limit}_{color_space}]$ denotes the range of intensity values of colour channels for enhanced rail lines in respective colour spaces.

Therefore, the ability of ACSM method to identify rail lines in DI captured in uneven sunlight intensity makes it adaptive.

5.1.1 Visual Analysis

The main aim of the visual analysis is to choose a colour space in which the respective colour space transformation is carried out on the track DI and the obtained railroad track mask consists of only rail lines. This way the railroad mask is able to eliminate uneven illumination, and background in DI. Here, track masks for D_2 and D_3 images in all four colour spaces are shown in Fig. 4.

Upon applying the thresholds, the RGB mask is created. It is observed in Fig. 4a that even though the rail lines in the RGB mask of D_2 are prominent along with clutter (ballast between the rail lines and along the left and right rail lines), the rail lines in the mask for D_3 are invisible. Hence, it is not feasible to consider an RGB colour space mask for further creation of edge map.

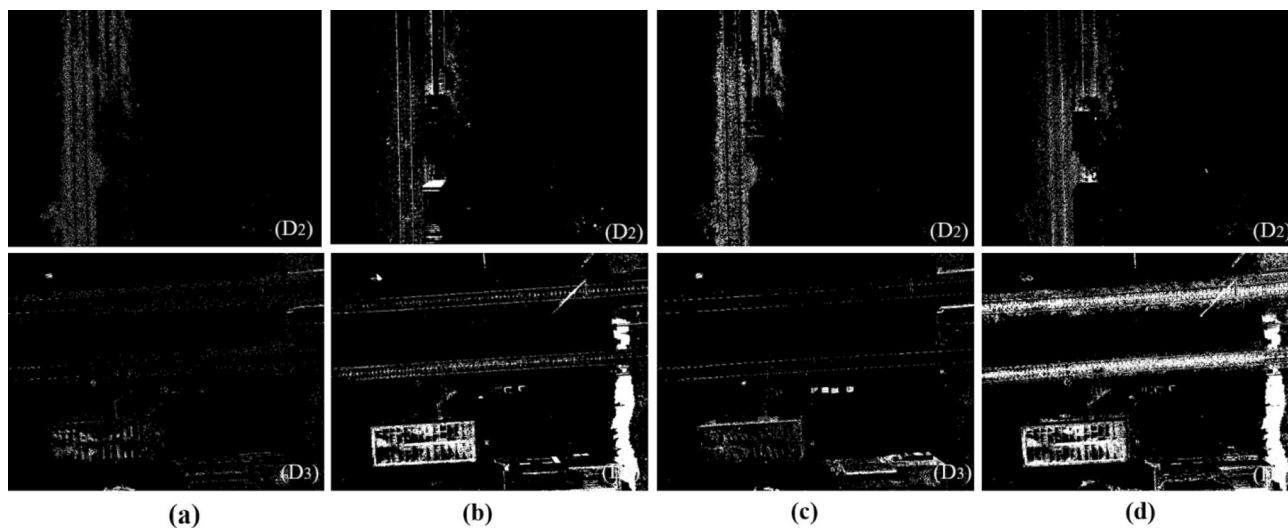


Fig. 4 Railroad Track masks for images D2 (above) and D3 (below) in **a** RGB Colour Space **b** HSV Colour Space **c** $L^*a^*b^*$ Colour Space **d** YCbCr Colour Space

Upon creation of HSV mask for D_2 , the rail lines can be definitively traced in the mask as shown in Fig. 4b. Also, the HSV mask for D_3 is present with quite distinct rail lines and negligent background clutter (ballast, pole lines, etc.) as seen in Fig. 4b. Hence, the HSV colour space has been considered for further evaluation.

Upon computation of the $L^*a^*b^*$ mask in Fig. 4c, we observe that the ballast surrounding the rail lines in D_2 mask is also distinctively visible. However, the rail lines in D_3 mask are still not visible in Fig. 4c. Hence, we discard this colour space due to inability to obtain proper railroad track mask comprising of only rail lines.

The YCbCr mask for D_2 in Fig. 4d shows the rail lines distinctly. However, the D_3 mask has white blobs and huge clutter as the railroad background, as shown in Fig. 4d. Hence, YCbCr colour space mask is not examined further.

Therefore, HSV colour space is selected for track mask creation. The masks for study images are shown in Figs. 6b, 7b, 8b, and 9b and they act as an input for edge map generation.

5.2 Edge Map Generation

The objective of this step is to generate an edge map from the railroad track mask (obtained in Sect. 5.1). For this purpose, a suitable edge detection algorithm needs to be selected and we have examined the following edge detection techniques: Prewitt, Sobel, Roberts, LOG and Canny [10, 23]. Upon evaluation of the comparison results, Canny is selected as the suitable edge detector [17] for rail lines.

The Canny edge detector poses the following advantages:

1. The Canny edge detector [46] uses hysteresis thresholding in which the connectivity characteristic is taken into consideration.
2. Also, in this detector, the computation of the first derivative (gradient) of Gaussian very closely approximates the operator, which causes optimization of the product of localization and signal-to-noise ratio.
3. Another advantage of Canny is that it uses the direction of the gradient to detect the rail edges.

The smoothing of the masked drone image is performed using the Gaussian kernel spread σ . The smoothed image is filtered with the Sobel kernel of size, Sob_{siz} , to compute the image gradients in the horizontal ($M_{i|x}$) and vertical ($M_{i|y}$) direction. The gradient magnitude ($M_{i|canny}$) and gradient direction (θ) are computed as in Eqs. (8) and (9):

$$M_{i|canny} = \sqrt{M_{i|x}^2 + M_{i|y}^2} \quad (8)$$

$$\theta = \tan^{-1} \left(\frac{M_{i|y}}{M_{i|x}} \right) \quad (9)$$

Then, hysteresis thresholding is applied on all the edges, with threshold values $minval$ and $maxval$, in order to remove small noise pixels (edge detections in the mask apart from rail line edges) and obtain localized rail line edge detections. These detections have been vividly outlined for all the study images using the Canny operator as shown in Figs. 6c, 7c, 8c, and 9c.

5.3 Geometrical Representation of Hough Parameter Space

The goal of this step is to take the generated edge map (computed in Sect. 5.2) as input, perform efficient Hough parameter space evaluation in order to obtain the geometrical representation of the rail lines and detect these lines at varying orientations. Thereafter, GSD-based line pair selection facilitates detection of rail line pairs at different flight heights. Railroad track extraction is then performed through coordinate transformation. This algorithm is termed as HT-GSD method and is described as Algorithm 2 in Fig. 5

5.3.1 Candidate rail line detection

In order to detect the rail lines, firstly their shape representation is obtained. To accomplish this task, we compute Hough transform [47] on the railroad track edge map for the detection of rail lines as straight lines that run through DI. A rail line is represented as a straight line in the Cartesian coordinate as in Eq. (10):

$$y = gx + c \quad (10)$$

where (x, y) is a point on the line, g is gradient and c is y -intercept. A rail line is represented in the Hough transform in the polar coordinate system, as a function of two parameters ρ and θ , as depicted in Eq. (11):

$$\rho = x \cos \theta + y \sin \theta \quad (11)$$

where, ρ = perpendicular distance from the coordinate origin $((0, 0)$, top-left corner in the image) to the line. θ is the angle formed by ρ . This geometrical representation of Hough parameter space in the DI is as shown in Figs. 6a, 7a, 8a and 9a. The Hough transform for the study images is computed in Figs. 6d, 7d, 8d, and 9d.

In order to hold the values of the two parameters, a representation of the Hough parameter space is created in the form of a 2-D array, $Acc(\rho, \theta)$. This array will store the (ρ, θ) value pairs of the detected lines. The values of ρ are a function of distance resolution, ρ_{acc} (in pixels), and values of θ are governed by angle resolution θ_{acc} (in radians). Here, ρ_{acc}

Algorithm 2: Hough Transform-Ground Sample Distance(HT-GSD)

```

Input: Hough parameter space  $houg(\rho_t, \theta_t)$ 
Output: all possible  $(L_1, L_2)$  rail line pairs
1. for  $t:=1:size(houg)$  :
2. if  $\rho_{first} > 0, \rho_t > 0$  or  $\rho_{first} < 0, \rho_t < 0$  :
     $\rho_i := \rho_t$ 
     $\theta_i := \theta_t$ 
3. return  $\rho_i, \theta_i$ 
4.  $\rhoLines(\rho_i, \theta_i)$ 
5. Set  $lin, lin_{sen}$ 
6. for  $i:=1:range(\rhoLines)$ 
7. if  $\rho_i - lin_{sen} \leq \rho_r \leq \rho_i + lin_{sen}$  :
8. return  $i+1$ 
     $\rho_s := \rho_i$ 
     $\theta_s := \theta_i$ 
9. else
     $\rho_{s+1} := \rho_r$ 
     $\theta_{s+1} := \theta_r$ 
10. return  $senLines(\rho_s, \theta_s)$ 
11. Set  $N_r$ 
12. for  $s:=1:range(senLines)$  :
13. if  $N_r - pix_{max} \leq (abs(\rho_{s+1}) - abs(\rho_s)) \leq N_r + pix_{max}$  :
     $\rho_i := \rho_s$ 
     $\theta_i := \theta_s$ 
     $\rho_{t+1} := \rho_{s+1}$ 
     $\theta_{t+1} := \theta_{s+1}$ 
14. return  $\rho_i, \theta_i$ 
15.  $\rhoLines(\rho_i, \theta_i)$ 
16. Set  $L_1 = (p1, p2), L_2 = (p3, p4), lextract$ 
17. if  $\rho_{first} > 0$  and  $\theta_{first} \in [1, 180]$  :
18. compute  $x1, x2, x3, x4$  from  $(\rho_i, \theta_i)$ 
19. return  $p1, p2, p3, p4$ 
20. else  $\rho_{first} < 0$  and  $\theta_{first} \in [0, 180]$  or  $\rho_{first} > 0$ 
    &  $\theta_{first} < [0-1]$  :
21. compute  $x1, x2, x3, x4$  from  $(\rho_i, \theta_i)$ 
22. return  $p1, p2, p3, p4$ 
23. return possible  $(L_1, L_2)$  rail line pairs
    
```

Fig. 5 HT-GSD algorithm

and θ_{acc} denote the minimum difference between any two ρ and θ values, respectively. The generated rail lines cover a range of angles $\theta \in [0, 180]$, and the angle is then measured in radians. If the angle resolution (θ_{acc}) is set to a larger value lesser (ρ, θ) pairs are recorded, however, this may cause the orientation precision of the rail lines to decrease. All drone images are of the same size but the length of the rail lines may differ as they are orientated at different angles (as seen in Fig. 1). The parameter values (ρ_t, θ_t) of Hough space with the number of votes equivalent or greater than the threshold H_{the} , which indicates minimum line length or the minimum number of intersections to detect a line, are recorded in $houg(\rho_t, \theta_t)$ for the corresponding lines to be selected as rail line candidates.

5.3.2 Parameter space evaluation

The candidate rail lines in DI (computed in Sect. 5.3.1) are aligned in horizontal or vertical direction at different orientations. The analysis of this rail line alignment (direction) is important so that all rail lines located horizontally or vertically at varying orientations are detected. In the geometry of rail lines, the Hough parameters ρ and θ play a fundamental role. Therefore, it is essential to extract parameter space values ρ_i and θ_i from $houg(\rho_t, \theta_t)$, which upon Cartesian Coordinate transformation produce lines that map approximately to the rail line locations in the right direction and at proper orientations in DI. It is observed that ρ_{first} , the first ρ value of $houg(\rho_t, \theta_t)$ is indicative of the direction of the rail lines in the DI and is examined to determine if the rail lines are horizontally or vertically oriented. For this, the following relationships are evaluated:

$$\rhoLines(\rho_i, \theta_i) = \begin{cases} \rho_i = \rho_t & \text{If } \rho_{first} > 0, \rho_t > 0 \text{ or} \\ \theta_i = \theta_t & \rho_{first} < 0, \rho_t < 0 \\ t = t + 1 & \text{Otherwise} \end{cases} \quad (12)$$

where, ρ_i = value of ρ in every iteration of the Hough parameter space $houg(\rho_t, \theta_t)$, θ_i = value of θ corresponding to ρ_i , ρ_i = value of ρ_t greater or lesser than zero based on conditions evaluated in Eq. (12) and θ_i = value of θ corresponding to ρ_i . These parameter values (ρ_i, θ_i) are recorded and updated in another array $\rhoLines(\rho_i, \theta_i)$. The conditions stated in Eq. (12) are evaluated in both the cases, i.e. if the orientation of all the rail lines is the same or if the orientation of rail lines differ.

After examination of the aforementioned conditions, now $\rhoLines(\rho_i, \theta_i)$ comprises of Hough parameter values, ρ_i and θ_i corresponding to the rail lines which would match to the rail line orientations and directions at their respective locations in DI. It is now important to determine another

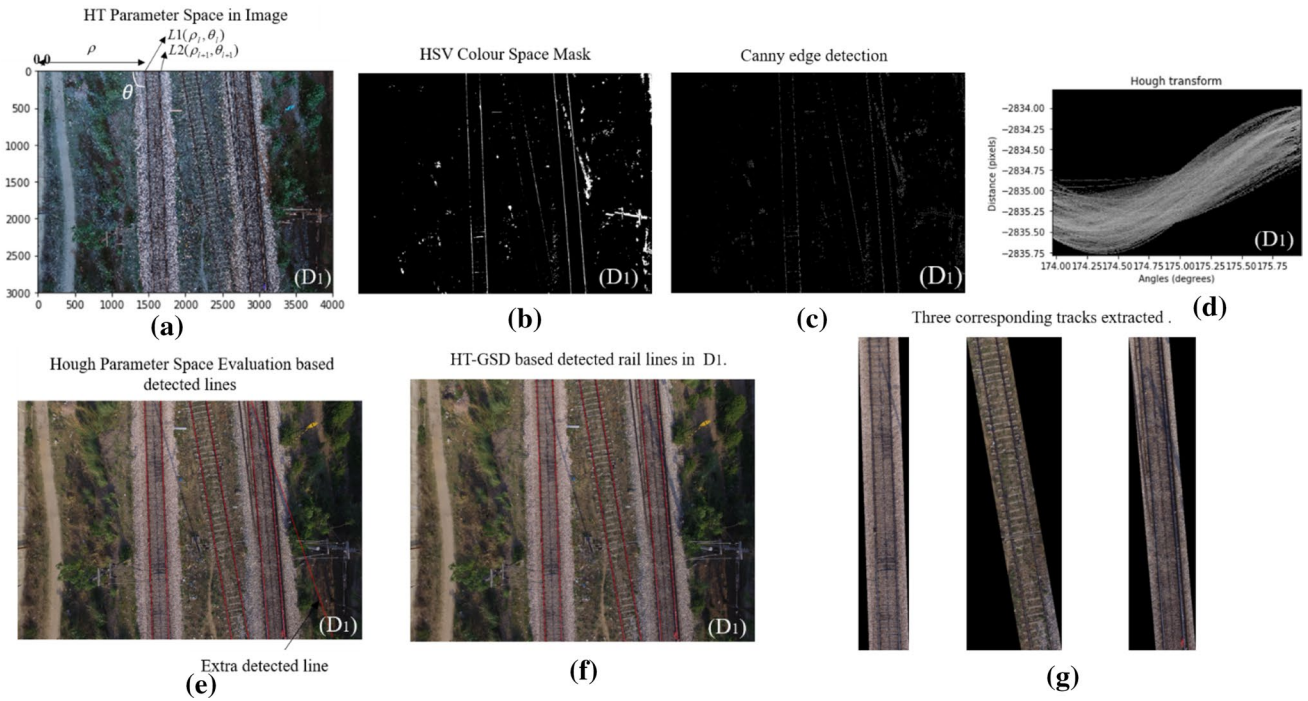


Fig. 6 Study Image 1: **D₁** **a** Geometrical Representation of Hough Parameter Space **b** HSV Colour Space Mask **c** Canny edge detection **d** Hough Transform **e** Detected lines based on Hough Parameter Space Evaluation **f** HT-GSD-based detected lines **g** Corresponding extracted tracks

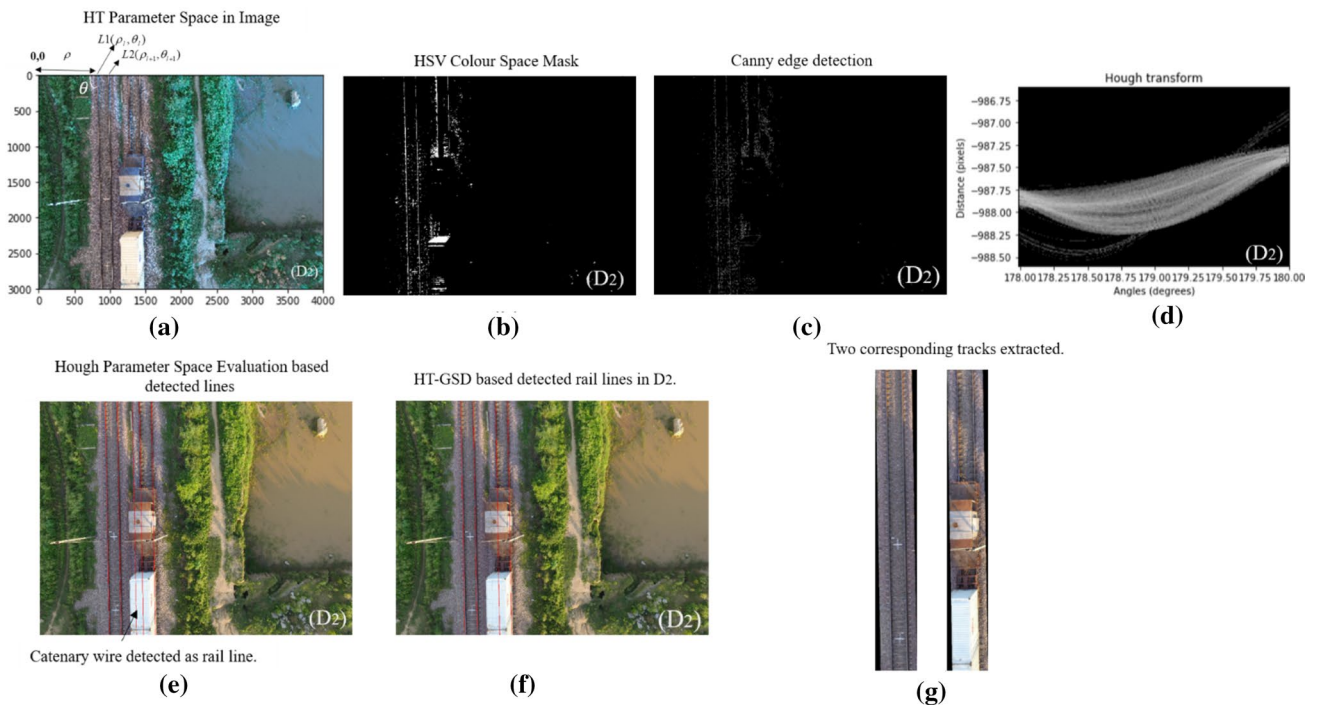


Fig. 7 Study Image 2: **D₂** **a** Geometrical Representation of Hough Parameter Space **b** HSV Colour Space Mask **c** Canny edge detection **d** Hough Transform **e** Detected lines based on Hough Parameter Space Evaluation **f** HT-GSD-based detected lines **g** Corresponding extracted tracks

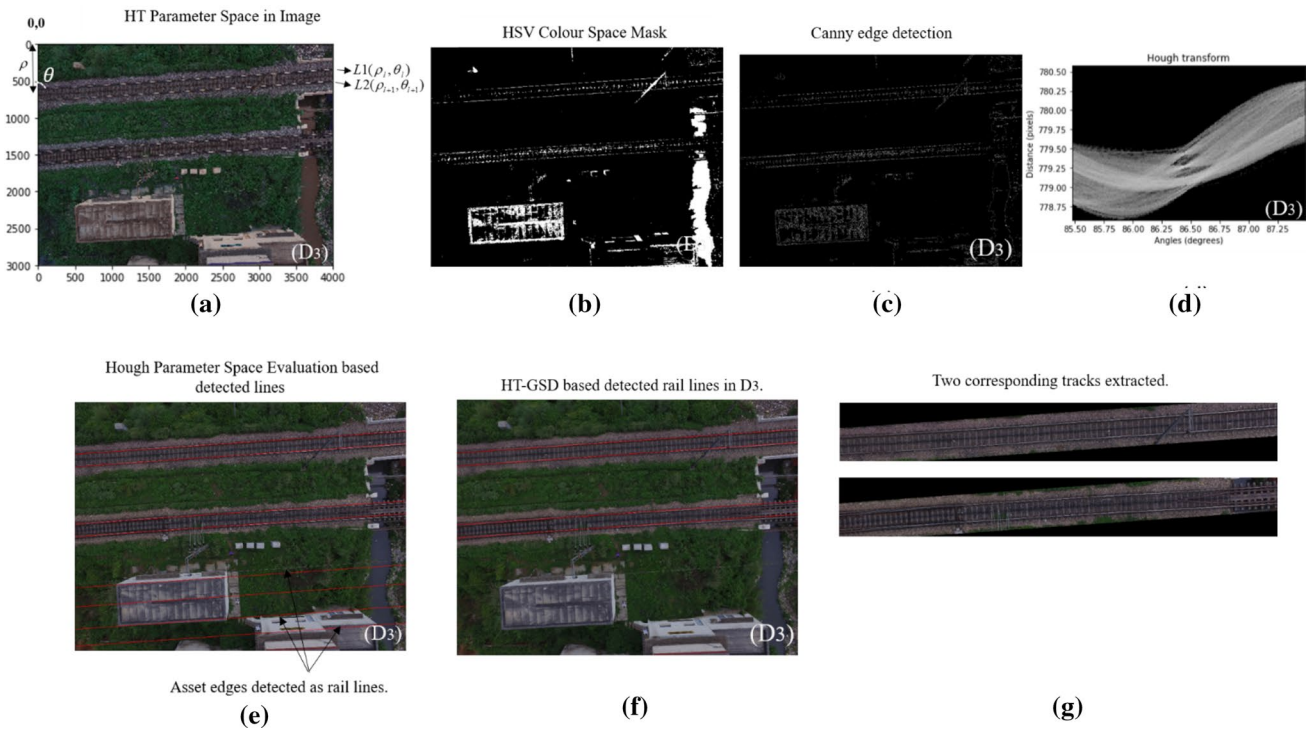


Fig. 8 Study Image 3: D_3 **a** Geometrical Representation of Hough Parameter Space **b** HSV Colour Space Mask **c** Canny edge detection **d** Hough Transform **e** Detected lines based on Hough Parameter Space Evaluation **f** HT-GSD-based detected lines **g** Corresponding extracted tracks

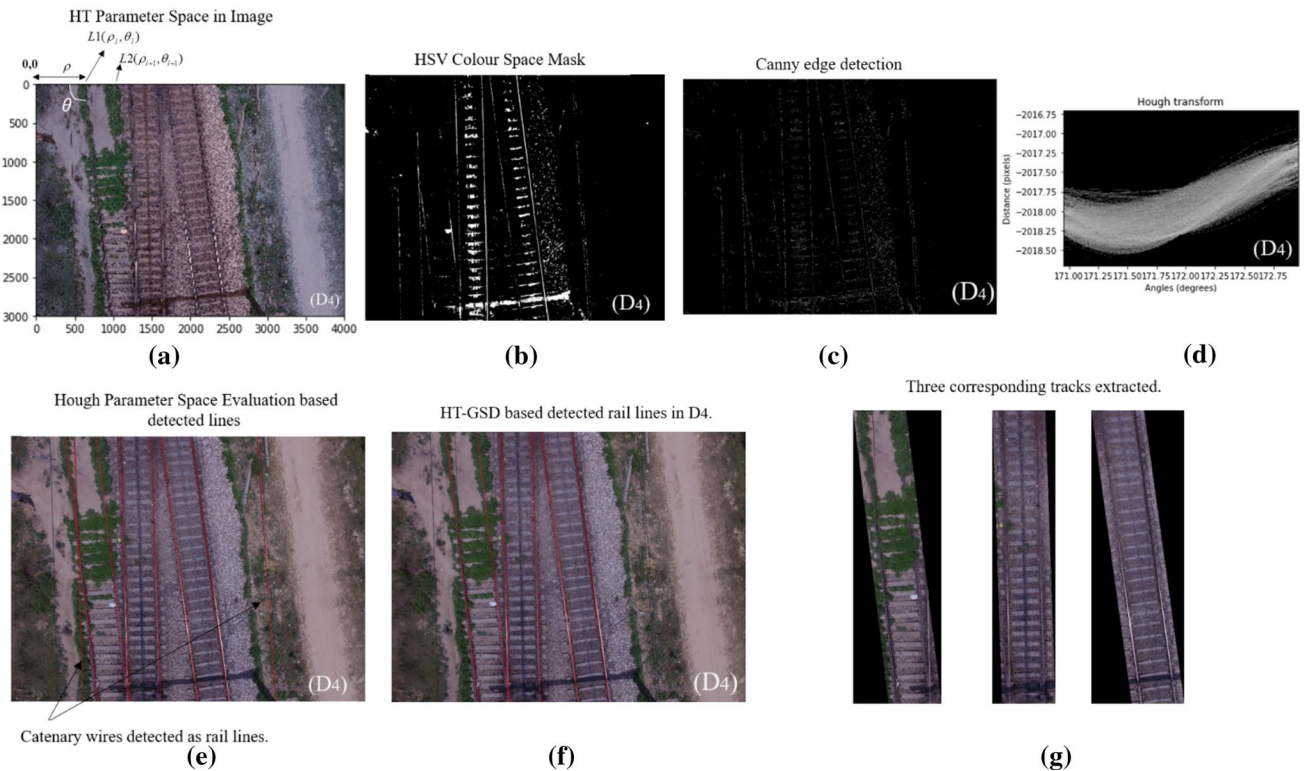


Fig. 9 Study Image 4: D_4 **a** Geometrical Representation of Hough Parameter Space **b** HSV Colour Space Mask **c** Canny edge detection **d** Hough Transform **e** Detected lines based on Hough Parameter Space Evaluation **f** HT-GSD-based detected lines **g** Corresponding extracted tracks

parameter lin which is the number of rail lines to be detected in the DI. Sometimes, setting a very small value for lin does not output all required rail lines while setting a very large value may provide with spurious output in the form of more number of lines than the actual number of rail lines present in DI. As much as it is essential to select the correct value of lin , it is important to remove these extra line detections except rail lines (if any). Therefore, in order to achieve this, the sensitivity range for lin is set to lin_{sen} pixels.

The detections within the sensitivity range are performed as follows:

$$senLines(\rho_s, \theta_s) = \begin{cases} i = i + 1 & \rho_i - lin_{sen} \leq \rho_r \leq \rho_i + lin_{sen} \\ \rho_s = \rho_i \\ \theta_s = \theta_i \\ \rho_{s+1} = \rho_r \\ \theta_{s+1} = \theta_r & \text{Otherwise} \end{cases} \tag{13}$$

where ρ_i, ρ_r denotes any pair of ρ parameter values in $\rhoLines(\rho_i, \theta_i)$. If any value ρ_r falls within the range given in Eq. (13), it implies the pair ρ_i, ρ_r values are placed within the lin_{sen} range and hence, ρ_r is discarded. Otherwise, both the ρ_i, ρ_r values are retained in the parameter space and the corresponding θ_i, θ_r are the orientation values of the rail lines obtained in $senLines(\rho_s, \theta_s)$. The detections are shown in Figs. 6e, 7e, 8e and 9e.

5.3.3 GSD-based rail line pair selection

In Sect. 5.3.2, even though we have tried to remove the spurious rail line detections, it is possible that such detections may still remain outside the range as defined in Eq. (13). Consequently, GSD-based rail line pair selection method is accurate for the selection of valid rail lines (of the rail line pair) and subsequent extraction of railroad track. GSD is calculated in Sect. 4.3 from Eqs. (1–3). The broad gauge length is represented by $track_{gauge}$. Consequently, the distance in pixels between the rail line pair, N_r , can be calculated as:

$$N_r = track_{gauge} / GSD \tag{14}$$

In this method, if the distance between the two rail lines is equal to N_r , they are parallel and belong to one rail line pair. To perform this computation, the following condition in Eq. 15 is computed between any pair of ρ values ρ_s, ρ_{s+1} in $senLines(\rho_s, \theta_s)$:

$$Lines(\rho_l, \theta_l) = \begin{cases} \rho_l = \rho_s & \text{If} \\ \theta_l = \theta_s & N_r - pix_{max} \leq (abs(\rho_{s+1}) - abs(\rho_s)) \\ \rho_{l+1} = \rho_{s+1} & \leq N_r + pix_{max} \\ \theta_{l+1} = \theta_{s+1} \\ s = s + 1 & \text{Otherwise} \end{cases} \tag{15}$$

where pix_{max} denotes the maximum number of pixels around the rail line which helps determine the range within which a rail line detection is considered as one of the rail lines of the rail line pair. The Hough space $Lines(\rho_l, \theta_l)$ gives the parameter values of the selected rail line pairs which satisfy the distance condition in Eq. (14) while calculating N_r and these line pairs are mapped to the exact rail line locations in the DI. The HT-GSD-based line pair selection has been depicted visually in Figs. 6f, 7f, 8f and 9f for different study images.

5.3.4 Coordinate system transformation-based railroad track extraction

Once the parameters for the rail line pair have been calculated, as in the previous Sect. 5.3.3, it is now essential to obtain the Cartesian line forms as in Eq. (10) of the rail lines (of the rail line pair) from the corresponding parametric forms shown in Eq. (11). A rail line pair is represented by two lines L_1 and L_2 , as shown in Figs. 6a, 7a, 8a and 9a. The rail line L_1 is described by points $(p1, p2)$ and another rail line L_2 is denoted by points $(p3, p4)$. For railroad track extraction, the coordinates of the rail line points need to be shifted by a value of $lextract$ on the left side of the left rail line and on the right side of the right rail line. The value of $lextract$ denotes the number of pixels (along the rail lines) which constitutes the ballast area along the left and right rail lines.

The coordinates are calculated by evaluating a (ρ_l, θ_l) pair under the following conditions:

1. If $\rho_{first} > 0$ and $\theta_{first} \in [1, 180]$:

$$x1 = \rho_l / \sin(\theta_l)$$

$$x2 = (\rho_l - I_w * \cos(\theta_l)) / \sin(\theta_l)$$

$$x3 = \rho_{l+1} / \sin(\theta_{l+1})$$

$$x4 = (\rho_{l+1} - I_w * \cos(\theta_{l+1})) / \sin(\theta_{l+1})$$

The coordinates of the two endpoints for the rail line L_1 , used for track extraction, are $p1(0, x1 - letract)$ and $p2(I_w, x2 - letract)$. Also, the endpoint coor-

dinates for rail line L_2 are $p3(0, x3 + l_{extract})$ and $p4(I_w, x4 + l_{extract})$.

2. If $\rho_{first} < 0$ and $\theta_{first} \in [0, 180]$ or $\rho_{first} > 0$ and $\theta_{first} \in [0 - 1]$:

$$x1 = \rho_l / \cos(\theta_l)$$

$$x2 = (\rho_l - I_h * \sin(\theta_l)) / \cos(\theta_l)$$

$$x3 = \rho_{l+1} / \cos(\theta_{l+1})$$

$$x4 = (\rho_{l+1} - I_h * \sin(\theta_{l+1})) / \cos(\theta_{l+1})$$

For track extraction in this scenario, the coordinates of the two endpoints for the line L_1 are $p1(x1 - l_{extract}, 0)$ and $p2(x2 - l_{extract}, I_h)$. The endpoint coordinates for line L_2 are $p3(x3 + l_{extract}, 0)$ and $p4(x4 + l_{extract}, I_h)$.

The above evaluated conditions calculate the coordinates of L_1 and L_2 endpoints in both the scenarios. Therefore, coordinate transformation facilitates railroad track extraction as shown in Figs. 6g, 7g, 8g and 9g.

6 Experimental results and discussions

6.1 Experimental setup

The track images have been captured by the DJI quadcopter. The sensor parameter values of the drone are $Flen = 3.61$ mm, $S_w = 6.16$ mm, $S_h = 4.62$ mm. The image dimensions are $I_h = 3000$ pixels and $I_w = 4000$ pixels. The proposed algorithm has been implemented as a software model using the OpenCV library function [48] and MATLAB R2016a. For evaluation, we have chosen 4–5 study images each (including the study images) from the acquired datasets for representation of uneven illumination, different rail line orientations and varied flight heights in different railroad environments.

6.2 Analysis of colour spaces for mask creation

An extensive evaluation of the colour space thresholds is performed to obtain a railroad track mask such that only rail lines are visible.

In order to obtain the railroad track mask in RGB colour space an extensive evaluation of thresholds has been carried out upon the two study images D_2 and D_3 , as shown in Fig. 4 a. The range for all three channels in RGB $\in [0-255]$. For thresholding, the $low_limit_{color_space}$ values for channel 1, channel 2 and channel 3 are set to 110, 14

and 119, respectively, and $upper_limit_{color_space}$ values are set to 158, 119 and 174, respectively.

In order to obtain the mask in HSV colour space, the hue range $\in [0, 179]$, saturation range $\in [0, 255]$ and value range $\in [0, 255]$ are marked in the study images as shown in Fig. 4b. For this purpose, the $low_limit_{color_space}$ and $upper_limit_{color_space}$ for mask thresholds for all three channels are set to [100, 40, 40] and [130, 255, 255], respectively.

$L^*a^*b^*$ expresses colour as: L^* for lightness which scales from 0 (black)-100(white), a^* axis and b^* axis both in the range $\in [-100, 100]$, which represent green (-) – red (+) and blue (-) – yellow (+) component, respectively. In order to obtain the masks, shown in Fig. 4c, the $low_limit_{color_space}$ and $upper_limit_{color_space}$ values for all three channels have been set to 43.419, 0.751, -31.492 and 63.950, 39.603, -5.247, respectively.

In the YCbCr, Y represents the luminance (intensity-based information) component and ranges from 16 to 235. The Cb, Cr channels are the chrominance (colour related) components both within the range $\in [16, 240]$. Also, the threshold values for the Y, Cb, Cr channels range from $low_limit_{color_space}$: 40, 130, 92, respectively, to $upper_limit_{color_space}$: 82, 194, 137, respectively, for obtaining the masks as shown in Fig. 4d.

6.3 Steps for Rail line detection and track extraction

1. In the ACSM method, it has been observed that rail lines have been distinctly identified upon segmentation only in the HSV colour space railroad track mask, as seen in Fig. 4b and discussed in Sect. 5.1. This colour space handles identification of rail lines well in non-uniform sunlight illumination.
2. Subsequently, edge detection is necessary in the HSV mask. After an exhaustive analysis of different sets of parameter values, it is observed that the following set of canny edge parameter values: Gaussian spread, $\sigma = 5 \times 5$, $Sob_{siz} = 3$, $minval = 50$, $maxval = 200$, projects sharp rail line edges as depicted in Figs. 6c, 7c, 8c and 9c. These edges are necessary as input for shape and geometric representations.
3. For computing shape representations of rail lines from edge map using Hough transform, the Hough parameter space resolutions are set to $\rho_{acc} = 2$ pixels, $\theta_{acc} = 0.017453$ radians (1°). The value for H_{the} is set to 100 pixels.
4. The values for number of lines $lin = 10, 18, 20, 25, 30$ lines have been chosen such that the value of lin is able to detect all the rail lines in all DI which may include spurious detections also.

5. The value for lin_{sen} is set to ± 80 pixels. The range is suitable in order to remove spurious rail line detections and is generally selected less than N_r . This is necessary to reduce the computational overhead for the next step by reducing the number of comparisons for line pair selections.
6. The GSD is calculated at any given flight height using Eqs. (1–3) and determines the pixel size on the ground as seen in Table 1. The broad gauge of tracks, $track_{gauge} = 1676$ mm (≈ 168 cm). The value of N_r is calculated for any given flight height using Eq. (14) which facilitates rail line pair selection. Two rail lines are classified as rail line pair when the distance (in pixels) between these two lines is evaluated as per Eq. (15). The value for $pix_{max} \in [5, 14]$. The $lextract$ value has been set to 100 pixels (the width in pixels) in order to obtain the left and right side of the rail lines for extraction of the railroad track.

The aforementioned empirically selected parameter values lead to correctly matched rail lines (in red colour). This facilitates respective automated railroad track

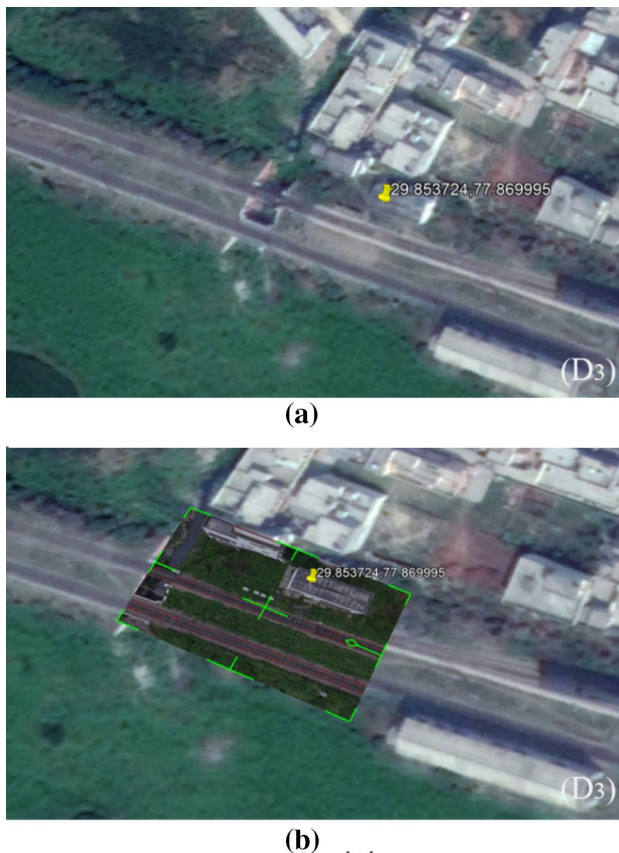


Fig. 10 a Google Earth Imagery for D_3 b Rail lines detected in red in image overlapped on Google Earth

extractions despite multi-rail lines and complex background as shown in Figs. 6g, 7g, 8g and 9g.

6.4 Evaluation and validation of the proposed method

In order to validate our proposed method, the values of Central Latitude: $29^{\circ}51'13.407399''N$ Central Longitude: $77^{\circ}52'11.9761''E$ for study image D_3 are extracted and marked over the Google Earth image of that area as seen in Fig. 10a. The study image with detected rail lines in red colour is then overlapped over the respective Google Earth image as seen in Fig. 10b. It has been observed that these detections have matched well even at locations with difficult terrain.

The efficiency of our proposed algorithm DroneRTEF(ACSM + HT-GSD) for drone images is evaluated using the following standard performance metrics: *Precision*, *Recall* and *Accuracy(ACC)* [49]. Before discussing these metrics, we should know about some prior notations. Suppose we classify our detections into two classes A (a rail line) and B (not a rail line), then true positive (TP) denotes number of correct matches (detections) which belong to A and are correctly identified as A, False Negative (FN) represents matches not correctly detected, False Positive (FP) denotes proposed matches that are incorrect and True Negative (TN) represents non-matches that are correctly rejected.

Precision is the measure of the proportion of the detected relevant instances or positive predictive value. It is represented as the ratio of TP and its sum with FP (depicted in Eq. 16).

Table 3 Performance measures for DroneRTEF (proposed approach)

No. of lines	Precision	Recall	Accuracy
10lines	0.88	0.6875	0.603774
18lines	0.735294	0.78125	0.877395
20lines	0.75	0.84375	0.902778
25lines	0.693333	0.8125	0.902235
30lines	0.658228	0.83871	0.913146

Table 4 Performance measures for ACSM + HT

No. of lines	Precision	Recall	Accuracy
10lines	0.471429	0.825	0.4125
18lines	0.321429	0.931034	0.313953
20lines	0.303571	0.934066	0.297203
25lines	0.282857	0.925234	0.276536
30lines	0.216667	0.938144	0.213615

Fig. 11 DroneRTEF versus ACSM+HT evaluation with **a** Precision **b** Recall **c** Accuracy

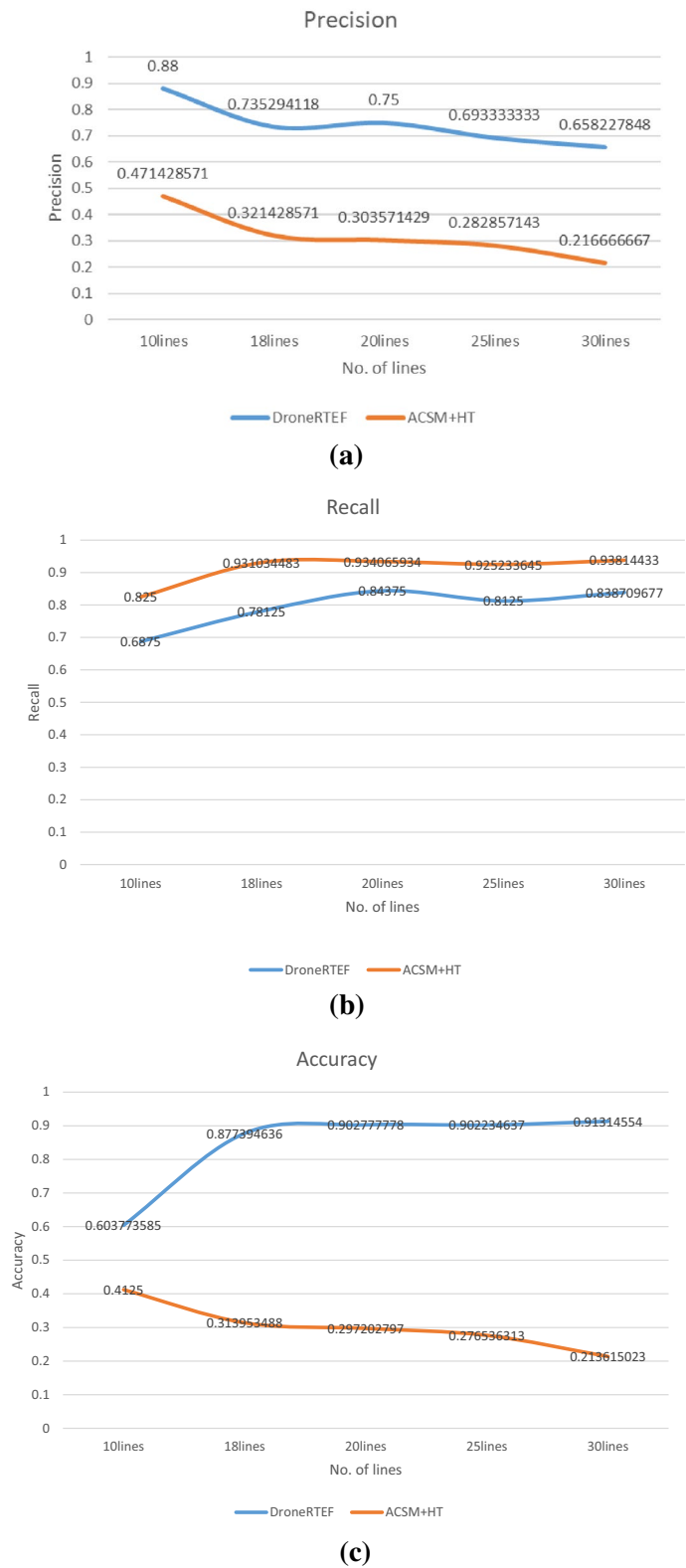


Table 5 Comparative Analysis

	Precision	Recall	Accuracy
K-NN mean Euclidean distances [29]	0.75	0.92	0.83
Inception V4 [30]	0.45	0.60	0.50
ACSM-HT(Computed)	0.30	0.934	0.29
DroneRTEF(Proposed Method)	0.75	0.843	0.902

$$\text{Precision} = \frac{TP}{TP + FP} \quad (16)$$

Recall, also called sensitivity, is the measure of the proportion of actual positives which are detected correctly. It is represented as the ratio of TP and its sum with FN (depicted in Eq. 17).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (17)$$

Accuracy (ACC) is the measure of correct detections made by the model over total number of detections(depicted in Eq. 18).

$$\text{Accuracy (ACC)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

In order to access the efficiency of our proposed approach, we have compared it with ACSM + HT on our newly acquired dataset. The metric analysis is performed on both the algorithms in order to calculate the difference in detections and show the effectiveness of our framework. The detection in ACSM + HT is considered after computing HT on the image. In our proposed approach, the detection is considered after parameter space evaluation. The evaluations on the metrics are performed with values of $lin = 10, 18, 20, 25, 30$ lines. The results are shown in Tables 3 and 4 for proposed approach and ACSM-HT approach, respectively.

The comparative analysis is obtained with three other methods K-NN mean Euclidean distances, Inception V4, ACSM-HT as shown in Table 5. The results have been obtained on our dataset. As observed the highest precision and accuracy values for the proposed framework DroneRTEF indicates robustness of the model upon comparison with other methods. However, recall values can be viewed to be comparable to K-NN and ACSM-HT method. The adaptive behaviour of the method can be inferred from the above results.

6.5 Discussion

The comparative analysis of Precision, Recall and Accuracy metrics between both the approaches is shown in Fig. 11a, b and c, respectively. These values have been computed for $lin = 10, 18, 20, 25, 30$ lines. In Fig. 11a, it has been observed that the Precision metric values for all lin values are more as

compared to the other method. Therefore, it can be inferred that the percentage of detected relevant lines is higher for the proposed approach as compared to the other method. In Fig. 11b, it has been observed that the recall values are higher for the other method. Recall values denote the proportion of actual positives identified correctly. In the proposed approach, it is observed that for each of the values in $lin = 10, 18, 20, 25, 30$ lines some of the actual detected lines are eliminated during the parameter evaluation method while removing the extra detections. Therefore, the recall value percentage is little lesser for our approach. The accuracy metric values are seen in Fig. 11c. It has been observed that the accuracy percentage is much higher for DroneRTEF. Therefore, it can be inferred that our proposed approach is efficient as compared to other algorithm for all the rail line values.

7 Conclusion and Future Work

In this paper, we have proposed DroneRTEF, a novel adaptive railroad track extraction framework for DI. This framework is divided into two stages. The identification of the rail lines is performed by ACSM. ACSM is a global image enhancement method in which colour space-based segmentation is performed for creation of a railroad track mask for DI comprising of only the rail lines. This approach facilitates rail line identification in DI with uneven illumination. Another step is Hough parameter space analysis-based novel HT-GSD algorithm. In this algorithm, hough parameters are evaluated in order to compute the shape and geometrical representation of lines in DI. This parameter space evaluation helps in the detection of identified rail lines at varying orientations, in different directions (horizontally or vertically) and also at different heights. This facilitates railroad track extraction. Our proposed framework is validated on large datasets and has achieved an accuracy of 90%. Therefore, the framework provides with the following advantages:

- 1 Drone is introduced as an IAS and provides with advantages such as image acquisition in inaccessible locations, is operative even during various railroad operations and provides cost-effectiveness.
- 2 A new dataset has been developed to take into consideration various railroad environment scenarios.
- 3 It is inferred from the experiments that the proposed method is adaptive to uneven illumination, varying rail line orientations and change in flight heights during track extraction in DI.
- 4 The algorithm has been tested and evaluated for large datasets in order to consider all railroad environment conditions and examine them for track extraction. This is in order to ensure safety of people travelling by railroad under any circumstances.

Appendix A

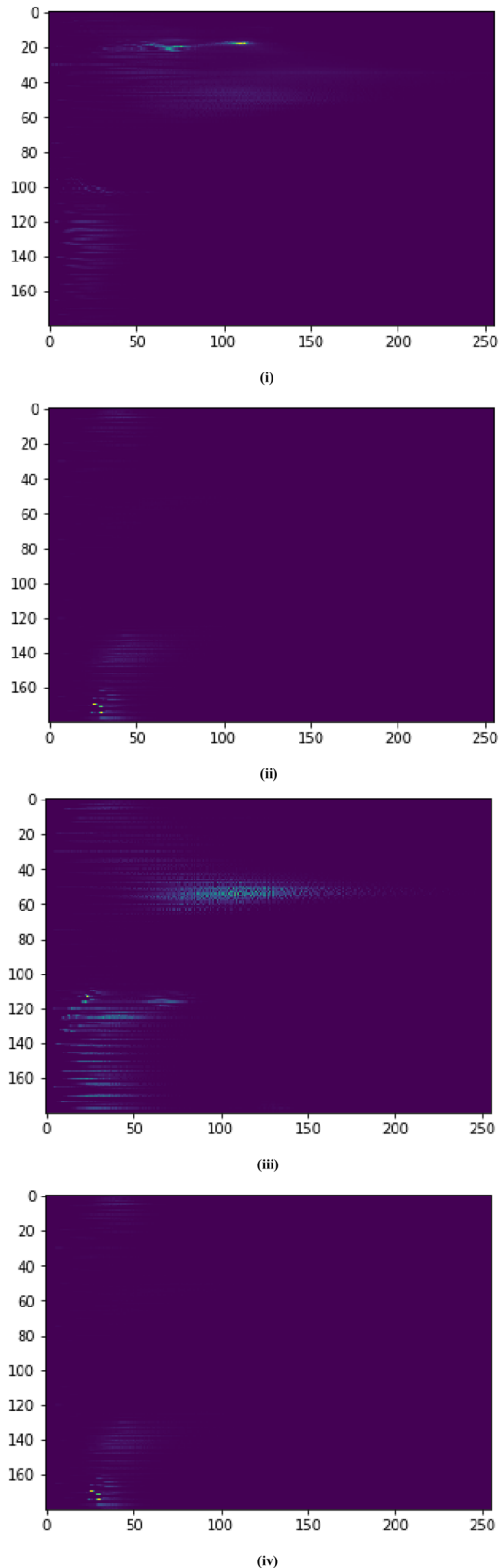


Fig. A1 HSV histograms for (i) D_3 (ii) D_4 (iii) D_2 (iv) D_1

- The effectiveness of the framework serves as an advantage for fast railroad inspections, obstacle identification and driver assistance systems.

In future, a technique to discuss removal of small recall values may be discussed. This is major importance as detection of rail lines is of paramount importance. Improvement of the HT-GSD method is proposed to be taken under consideration.

Acknowledgements The authors are thankful to RailTel, India for supporting this work.

References

- Min Y, Xiao B, Dang J, Yue B, Cheng T (2018) Real time detection system for rail surface defects based on machine vision. *EURASIP J Image Video Process* 2018(1):3.
- Kaleli F, Akgul YS (2009) Vision-based railroad track extraction using dynamic programming. In: 2009 12th International IEEE Conference on Intelligent Transportation Systems, (IEEE, 2009), pp 1–6
- Kim Z, Cohn TE (2004) Pseudoreal-time activity detection for railroad grade-crossing safety. *IEEE Trans Intell Transportation Syst* 5(4):319–324.
- Feng H, Jiang Z, Xie F, Yang P, Shi J, Chen L (2013) Automatic fastener classification and defect detection in vision-based railway inspection systems. *IEEE Trans Instrumentation Meas* 63(4):877–888.
- Espino JC, Stanculescu B (2012) Rail extraction technique using gradient information and a priori shape model. In: 2012 15th International IEEE Conference on Intelligent Transportation Systems, (IEEE, 2012), pp. 1132–1136
- Espino JC, Stanculescu B, Forin P (2013) Rail and turnout detection using gradient information and template matching. In: *Intelligent Rail Transportation (ICIRT)*, 2013 IEEE International Conference on (IEEE, 2013), pp 233–238
- Ashish, J., Wang, J., Yang, X., *et al.*: ‘TrackNet-A Deep Learning Based Fault Detection for Railway Track Inspection’ 2018.
- Qi Z, Tian Y, Shi Y (2013) Efficient railway tracks detection and turnouts recognition method using HOG features. *Neural Comput Appl* 23(1):245–254.
- Li Q, Ren S (2012) A visual detection system for rail surface defects. *IEEE Trans Syst Man Cybern C (Applications and Reviews)* 42(6):1531–1542
- Wang Z, Cai B, Chunxiao J et al (2016) Geometry constraints-based visual rail track extraction. In: 2016 12th World Congress on Intelligent Control and Automation (WCICA), (IEEE, 2016), pp 993–998
- Li Q, Ren S (2012) A real-time visual inspection system for discrete surface defects of rail heads. *IEEE Trans Instrumentation Meas* 61(8):2189–2199.
- Wu H, Siu W-C (2015) Real time railway extraction by angle alignment measure. In: 2015 IEEE International Conference on Image Processing (ICIP) (IEEE, 2015), pp 4560–4564
- Warsop T, Singh S (2012) A framework for computing asset-sighting distance. *Pattern Anal Appl* 15(4):427–444.

14. Nyberg RG, Gupta NK, Yella S, Dougherty MS (2014) Machine vision for condition monitoring vegetation on railway embankments, 2014.
15. Higgins C, Liu X (2018) Modeling of track geometry degradation and decisions on safety and maintenance: a literature review and possible future research directions. *Proc Inst Mech Eng F J Rail Rapid Transit* 232(5):1385–1397
16. Leizer GKK, Tokody D (2017) Radiofrequency identification by using Drones in railway accidents and disaster situations. *Interdisciplinary Description Complex Syst* 15(2):114.
17. Máthé K, Buşoniu L (2015) Vision and control for UAVs: a survey of general methods and of inexpensive platforms for infrastructure inspection. *Sensors* 15(7):14887–14916.
18. Jang J, Shin M, Lim S, Park J, Kim J, Paik J (2019) Intelligent image-based railway inspection system using deep learning-based object detection and Weber contrast-based image comparison. *Sensors* 19(21):4738.
19. Wu Y, Qin Y, Wang Z, Jia L (2018) A UAV-based visual inspection method for rail surface defects. *Appl Sci* 8(7):1028.
20. Smith EM (2016) A collection of computer vision algorithms capable of detecting linear infrastructure for the purpose of UAV control. PhD Thesis, Virginia Tech
21. Wei X, Wei D, Suo D, Jia L, Li Y (2020) Multi-target defect identification for railway track line based on image processing and improved YOLOv3 Model. *IEEE Access*, 2020.
22. Frauenthal JM (2015) Design and exploration of a computer vision based unmanned aerial vehicle for railroad health applications. PhD Thesis, Virginia Tech
23. Taştımur C, Karaköse M, Akın E (2016) A vision based condition monitoring approach for Rail Switch and level crossing using hierarchical SVM in railways. *Int J Appl Math Electron Comput* 2016, (Special Issue-1):319–325.
24. Teng Z, Liu F, Zhang B (2016) Visual railway detection by superpixel based intracellular decisions. *Multimedia Tools Appl* 75(5):2473–2486.
25. Arun K, Anushka S, Ankush A, Dharmendra S (2017) Vision based rail track extraction and monitoring through drone imagery (KICS, 2017)
26. Jarrett C, Perry K, Stol KA (2015) Controller comparisons for autonomous railway following with a fixed-wing UAV. In: 2015 6th International Conference on Automation, Robotics and Applications (ICARA) (IEEE, 2015), pp 104–109
27. Goldenshluger A, Zeevi A (2004) The Hough transform estimator. *Ann Stat* 32(5):1908–1932.
28. Le Saux B, Beaupère A, Boulch A, Brossard J, Manier A, Villemain G (2018) Railway detection: from filtering to segmentation networks. In: 'IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, (IEEE, 2018), pp 4819–4822
29. Banić M, Miltenović A, Pavlović M, Ćirić I (2019) Intelligent machine vision based railway infrastructure inspection and monitoring using UAV. *Facta Universitatis Series: Mech Eng* 17(3):357–364.
30. Ikshwaku S, Srinivasan A, Varghese A, Gubbi J (2019) Railway corridor monitoring using deep drone vision. In: *Computational Intelligence: Theories, Applications and Future Directions*, vol II. Springer, Cham, pp 361–372
31. Purica AI, Pesquet-Popescu B, Dufaux F (2017) A railroad detection algorithm for infrastructure surveillance using enduring airborne systems. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), (IEEE, 2017), pp 2187–2191
32. Hassanein AS, Mohammad S, Sameer M, Ragab ME (2015) A survey on Hough transform, theory, techniques and applications. arXiv preprint [arXiv:1502.02160](https://arxiv.org/abs/1502.02160).
33. Páli E, Mathe K, Tamas L, Buşoniu L (2014) Railway track following with the AR. Drone using vanishing point detection. In: 2014 IEEE international conference on automation, quality and testing, Robotics' (IEEE, 2014), pp 1–6
34. Wang W, Tan-Torres A, Hamledari H (2017) Lecture# 06: edge detection, computer vision: foundations and applications (CS 131), Stanford University
35. Ruzgienė B, Berteška T, Gečyte S, Jakubauskienė E, Aksamitauskas VČ (2015) The surface modelling based on UAV Photogrammetry and qualitative estimation. *Measurement* 73:619–627.
36. Otsu N (1979) A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 9(1):62–66.
37. Wong AK, Sahoo PK (1989) A gray-level threshold selection method based on maximum entropy principle. *IEEE Trans Syst Man Cybern* 19(4):866–871
38. Shafarenko L, Petrou H, Kittler J (1998) Histogram-based segmentation in a perceptually uniform color space. *IEEE Trans Image Process* 7(9):1354–1358.
39. Dev S, Lee YH, Winkler S (2017) Color-based segmentation of sky/cloud images from ground-based cameras. *IEEE J Selected Topics Appl Earth Observations Remote Sensing* 10(1):231–242.
40. Danelljan M, Shahbaz Khan F, Felsberg M, Van de Weijer J (2014) Adaptive color attributes for real-time visual tracking. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp 1090–1097
41. Shaik KB, Ganesan P, Kalist V, Sathish BS, Jenitha JMM (2015) Comparative study of skin color detection and segmentation in HSV and YCbCr color space. *Proc Comput Sci* 57:41–48.
42. Ford A, Roberts A (1998) *Colour space conversions*. Westminster University, London, pp 1–31
43. Poynton CA (1996) *A technical introduction to digital video*. Wiley, New York
44. Zheng S, Chai X, An X, Li L (2012) Railway track gauge inspection method based on computer vision. In: 2012 IEEE International conference on mechatronics and automation (IEEE, 2012), pp 1292–1296
45. Kotoulas L, Andreadis I (2003) 'Colour histogram content-based image retrieval and hardware implementation. *IEE Proc-Circuits Devices Syst* 150(5):387–393
46. John C (1986) A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intekkgence* 8(6).
47. Hough PV (1962) Patent and Trademark Office. U.S. Patent No. 3,069,654. Washington, DC: U.S
48. Bradski G, Kaehler A (2008) *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc., Sebastopol
49. Szeliski R (2010) *Computer vision: algorithms and applications*. Springer Science & Business Media

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.