

Successive Texture and Shape based Active Contours for Train Bogie Part Segmentation in Rolling Stock Videos

Kaja Krishnamohan¹, Ch.Raghava Prasad², P.V.V.Kishore³

Research Scholar, Department of ECE, Koneru Lakshmaiah Education Foundation, Guntur (DT), Andhra Pradesh, INDIA.¹
Department of ECE, Koneru Lakshmaiah Education Foundation, Guntur (DT), Andhra Pradesh, INDIA.^{2,3}

Abstract—Train Rolling Stock Examination (TRSE) is a procedure for checking damages in the undercarriage of a moving train at 30kmph. The undercarriage of a train is called bogie according to railway manuals. Traditionally, TRSE is performed manually by set of highly skilled personnel of the railway near to the train stations. This paper presents a new method to segment the TRSE bogie parts which can assist trained railway personnel for better performance and consequently reduce train accidents. This work uses visualization techniques as a pair of virtual eyes to help checking of each bogie part remotely using high speed video data. Our previous AC models are being supervised by a weak shape image which has shown to improve segmentation accuracies on a closely packed inhomogeneous train bogie object space. However, the inner texture of the objects in the bogies is found to be necessary for better object segmentation. Here, this paper proposes an algorithm for bogie parts segmentation as successive texture and shape-based AC model (STSAC). In this direction, texture of the bogie part is applied serially before the shape to guide the contour towards the desired object of interest. This contrasts with the previous approaches where texture is applied to extract object shape, losing texture information completely in the output image. To test the proposed method for their ability in extracting objects from videos captured under ambient conditions, the train rolling stock video database is built with 5 videos. In contrast to previous models the proposed method has produced shape rich texture objects through contour evolution performed sequentially.

Keywords—Automation of Train Rolling Stock Examination; level sets; shape priors; texture priors; export system models

I. INTRODUCTION

Visual automated testing of machines by computer algorithms has been gaining momentum in the past few decades. This increase can be attributed to factors such as high-resolution visual sensors, high speed cameras and more significantly the higher processing power of computers. Progressively, these advancements can be noticed in manufacturing industries, where the assembly lines are monitored visually by high speed cameras to identify defects in products manufacturing processes and packaging. Consequently, the manufacturing industry was revolutionized by visual monitoring technologies thereby improving productivity and quality of production. The long-term dependencies were higher revenues and lowered labour costs. Steadily visual automation has become industry's biggest challenge in promising new solutions to multitude of problems. One such problem that hadn't been explored was Train rolling stock examination.

Train Rolling Stock Examination (TRSE) is a budgeted

system on the Indian Railways operational space. The TRSE is currently being executed at every major train station across the Indian subcontinent and the world over to man the safety of passenger trains. Trains on Indian subcontinent carry around 10 million passengers per day. This has been the primary mode of commercially affordable long distance transportation on the planet. Safety of the train during transit is the most significant factor for the rail companies around the world and the foremost job for Indian Railways. Considering the number of train accidents from the past decades, the train transportation has been one of the safest mode of travel and is mostly attributed to rolling stock examination personal.

Train Rolling Stock Examination (TRSE) is a procedure for checking damages in the undercarriage of a moving train at 30kmph. The undercarriage of a train is called bogie according to railway manuals. The bogie consists of dynamic machinery on which the passenger car moves. It is made of wheels, break units, suspension, holding rods, springs, axle box, etc. There are around one hundred components in the bogie that cater for the train movement. The bogie parts have to be constantly monitored during transit as there go through extremities of pressures. The pressure on the bogie parts come due to inter part stroking between them during high speed motion of the train. This causes wear and tare in the bogie parts, which if not checked in time have caused extensive damage to the train causing derailment and human loss. To periodically check the bogie parts during transit, the long lasting and most trusted process is train rolling stock examination.

Traditionally, TRSE is performed manually by set of highly skilled personnel of the railway near to the train stations. Fig. 1 shows a rolling pit with trained railway employees noting the results of their examination (not in frame). These personal are trained for years to use their visual and auditory sensors to identify weaknesses in the bogie parts that can potentially cause an accident. Consequently, the noted risk factors are relayed to the nearest station maintenance crew for necessary repairs. Though the process of TRSE is full proofed, the system in the past hasn't been successful in preventing accidents and loss of life. The fact that the system is heavily dependent on human performance in naturalistic environments, which is dynamic in lighting, temperature, winds and water. Finally, it also depends on human emotional health at the time of the hour.

The goal of any railway company is to provide a safe transportation system. Despite their committed efforts through



Fig. 1. Manual Train Rolling Stock Examination at an Indian Railway Station.

times there are accidents, and many are during train movement. Manual TRSE needs an extra degree of support to perform without glitches. Technology started providing solutions to this age-old problem only in last two decades. Despite some conformations using sensors, there were no real solutions on the visual frontiers. Train Rolling Stock Examination (TRSE) is a conditional health monitoring system for damage detection in moving passenger bogies used to prevent trains from derailment. Currently, TRSE has been performed manually by experts through audio visual inspection on a moving train at both ends of a train station.

The commercially available sensors and signal processing technology used for TRSE has successfully identified only 20% of the total detect causing derailments. The large set of anomalies found during TRSE are predominantly visual in nature. Thus, a visual automation solution using computer vision algorithms can make the process more robust in preventing train accidents. To this end, the primary task is to extract bogie objects from a video sequence of moving train.

Hence, this chapter proposes a novel segmentation model on the videos of train bogies using a serial texture and shape-imposed level set evolution. The present generation of level set models use texture and shape priors for segmenting objects in an image. However, these models regularize the shape of the contour using the texture in the shape region resulting in a boundary shaped object with no texture. This limits the algorithms capacity to handle segmentation in real time video applications. To preserve texture during shape segmentation, we propose a novel serial texture and shape prior level set evolution model. We first present a theoretical framework for the model with various parameters for regularized contour evolution.

To demonstrate the applicability of the proposed method, experimentation and analysis is performed on benchmark image data and the real-time application, TRSE. The consistency of the algorithm is validated against the state – of – the – art level sets on TRS high speed video datasets. The results show that the proposed method is feasible in practice for segmenting texture preserving shapes in real time videos. The rest of the paper is arranged as follows. The second section gives the background motivation highlighting various gaps in current methods. Methodology and experimentation are provided in sections three and four, respectively. Final section of this paper

draws conclusions on the proposed method.

II. RELATED BACKGROUND

The promising and motivational research that inspired the formation of this thesis was industrial imaging solutions [1]. Industrial computer vision applications included a variety of image acquisition systems that use the captured videos to detect patterns during the product assembly. The most widely used are CMOS image sensors and hyperspectral sensors [2]. These sensors along with the embedded software has shown to be a valuable asset in bottling and beverage cans quality testing and discarding the faulty bottles on a high-speed assembly line [3].

The automobile industry and its robots use computer vision systems for wheel alignment to mirror inspections [4]. Largely, the operations performed by the software programs are designed to process the image of the object in question to make decisions on its quality and maintenance. The image processing methods used range from simple edge detection to as complex as filtering in frequency domain [5]. Consequently, offline testing of industrial products is on the rise from the last few decades due the availability of commercially viable sensors [6].

Subsequently, vision-based models have shown to provide accuracies on the higher side in most of the real-world industrial applications. This has motivated us to take up the study and investigate the problem to make the manual system of TRSE into an assistant for railway personal. The next session gives an insight into the current operating models of TRSE that are being prototyped and undergoing testing in railways across the world.

The railways around the world and the IR have adopted technologies for locomotive design, coaches, signalling systems, accident prevention with GPS, track maintenance [7], [8], [9], [10], [11]. The quintessential component would be maintenance of train coaches and bogies that are most likely to get damaged during their high-speed movement.

IR uses a highly trained human workforce to do the maintenance checks under the banner called Rolling Stock Examination (RSE) [12], [13]. The checking of the train happens during the train movement at less than 30KMPH near to the railway stations. The check log that is prepared during each RSE involves the visual analysis of the train undercarriage during running is accessible at [14].

The first models that were developed involved sensors along the track that measure parameters such as temperature, pressure, break wear and tear, acceleration etc along with a normal camera module [15]. This prototype is currently being tested in the code name KRATES - Konkan Railway Automated Train Examination System [16].

In this system, the objective of the camera was to have a visual examination at a remote location and no algorithms were proposed to automate the visual information. Apart from this, the camera is an RGB video camera with a frame rate of 30fps, which gives blurry images of bogie for automated processing.

However, the RSE involves checks for hanging parts, lose couplings in bogie parts, break bindings, broken components,

hot axle boxes and flat tyres. The following disadvantages were listed for manual TRSE which are limited to, Human factors – Biased Judgment, Heavy Workload – Wrong Judgment, Communication Lags – Between RSE and Maintenance dept., Ambient Nature – Weather Dependent and Commercially Draining.

A complete automated TRSE is quite possible with a large sensor network placed along the tracks. In the current scenario it is still a long-term plan for most of the rail networks due to issues like technology development, deployment and commerce. Despite these issues, this problem is quite challenging and an assistance to manual TRSE is proposed in through this work.

Narayanaswami [17] helps to unfold the connection between automation technologies in transportation to prevent accidents. Inspired by the ideas in [17], this work applies machine vision algorithms for discovering train undercarriage parts from videos of recorded. There are only few works on train safety research with computer vision, targeted at monitoring of rails, ballast and a few on rolling stock.

Sabato et al. [18] proposed methods to inspect train tyres as well as ballast using algorithms developed on 3D digital image correlation (DIC) calculations. It's a dual camera system with a marked displacement installed on a train running at 60Kmph that builds rails as a 3D image. The 3D DIC along with pattern projection models were applied to identify deformation of railway tracks.

Hart, et al. [19] extracted bogie parts for inspection using multispectral imaging camera sensors. The proposed dual camera model recorded multi spectral data as RGB (Red, Green and Blue) along with a thermal sensor to capture a running train which are further treated as panoramic view models.

The computer vision algorithm in [19] has been designed to spot elevated hot bogie parts such as wheel joints, Axle box, brake shoes, air conditioning blowers. Nevertheless, this system has been successful in identifying defective regions, but the motion blur in the video data poses a challenge to distinguish cold parts from hot objects.

Kim et al. [20], has proposed a curve fitting view to the problem of automated train break examination using image processing. The developed techniques use a trench hole establishment under the tracks to capture the break panels of a moving train. The method uses a fitting curve on the recorded images to progressively train the system to identify brake alignment attributes. Despite its excellent performance in real time, the setup cost creates a bottleneck for actual implementation.

The US patent from Sanchez, et al. [21], applies artificial vision for monitoring rolling stock using cameras mounted on the train. Currently, high speed trains such as TGV and bullet train use camera mounts to manually monitor the trains movements. However, videos captured using a camera system onboard a train is bound to induce numerous noises into the video data.

Kazanskiy and Popov [22], introduced a framework to integrate a lighting structure with anti-glare to record high contrast undercarriage videos which is further compressed for

quick processing to discover trains on tracks for monitoring rolling stock. This method gave a recipe for automating rolling stock in real time, notwithstanding the procedure for bogie object extraction.

Freid [23], provided an experimental setup under the train with lights focused on the bogie which is captured with a video camera. The work develops an algorithm using straightforward edge recognition techniques for isolating axle box and analyzing its heating profile by using thermal cameras. This model gives an understanding of the TRSE problem for automation and the need for research.

In [24] and [25], the authors offered a 3D reconstruction of the bogie parts for monitoring rail wheel surfaces and contact strips. The methods show effectiveness in identifying surface defects using 3D models by perfectly reconstructing moving parts. However, they are computationally inefficient for processing in real time. Further, it shows the difficulty in modelling defective surfaces for every possible problem beforehand in 3D.

The literature illustrates relatively small number of computer vision state-of-the-art algorithms that are being researched for TRSE. Moreover, the models from literature are inefficient to incorporate the TRSE process for micro level examination of bogie parts specifically. The goal of remotely monitoring system for TRSE is to identify defective and non-working parts that can be repaired timely to prevent mishaps. Earlier proposed frameworks offer very little inclination for research towards remote monitoring of TRSE. Hence, this paper proposes a new orientation to the TRSE with solutions to assist trained railway personnel for better performance. This work uses visualization techniques as a pair of virtual eyes to help checking of each bogie part remotely.

TRSE with video data has been attempted previously [26], [27], [28], [29] using active contours with shape prior models. The performance reported by these models were exceptionally good in terms of segmentation accuracy. However, these models were limited by their ability to provide the required accuracies due to homogeneous nature of pixels in the video sequence.

This gap in segmentation accuracy has been improved by applying local texture information around the object of convergence in the objective function defining the active contours. Moreover, the methods were derived from Chan Vese active contour models which will be discussed exclusively in the following sections. Traditionally, texture and shape based active contours use texture information in a region for shape segmentation [30], [31], [32], [33], [34]. In contrast, this work proposes to segment shape rich texture objects through contour evolution performed sequentially on a shape prior model.

This work proposes to segment video objects using the texture and shape based active contour models successively. The novelty lies in extracting shape and texture of the bogie parts accurately. Contrasting to the preceding results in [26] or correlated works [35], [36], [37], the proposed model in this paper will serially supervise texture and shape segmentation of TRSE video objects with inhomogeneity.

III. METHODOLOGY

This section starts by presenting a brief introduction about Chan Vese active contour models [38]. Further, it provides detailed methodology about the proposed successive texture and shape-based model for bogie part video segmentation.

A. Introduction to Chan Vese Active Contours

Active contours or level sets as they are mathematically named are a set of curves in image space that propagate towards the object edges defined by the image gradients. The most popular class of ACs are Chan Vese (CV) model without edges [38], which are described in the following section.

CV active contours evaluate to find a contour C in the image space $f(x) \forall x \in (R, R)$ through the energy function modelled as

$$E_{cv} = \omega_1 \int_C ds + \nu \iint_{C^I} C(x) dx + \omega_2 \left[\frac{1}{2} \iint_{C^I} (f(x) - \Phi^{(I)})^2 dx + \frac{1}{2} \iint_{C^E} (f(x) - \Phi^{(E)})^2 dx \right] \quad (1)$$

Where E_{cv} gives the energy function of the CV active contour. The first two terms in eq (1) are parameters controlling the contour's length and area with $\omega_1 > 0, \omega_2 > 0$ and $\nu > 0$. The last bracketed components try to adapt the model $C(x)$ to the image. The energy function in eq (1) is minimized iteratively to find the object boundaries through an initial contour $f : C \rightarrow R^2$ to $\Phi^{(I)}$ and $\Phi^{(E)}$. Where $\Phi^{(I)}$ and $\Phi^{(E)}$ define the inside and outside portions to Φ .

Eq (1) has been modified by applying the level set models in [38] as

$$E_{cv}(C, \Phi^{(I)}, \Phi^{(E)}) = \min_{C, \Phi^{(I)}, \Phi^{(E)}} \omega_2 \left[\iint_{C^I} (f(x) - \Phi^{(I)})^2 M(C(x)) + \iint_{C^E} (f(x) - \Phi^{(E)})^2 (1 - M(C(x))) dx \right] + \omega_1 \int_C |\nabla M(C(x))| dx \quad (2)$$

Where, $M_\varepsilon(C) = \frac{1}{2} \left[1 + \frac{2}{\pi} \tan^{-1} \left(\frac{C}{\varepsilon} \right) \right]$, $C \in \mathbb{R}$ gives the Heaviside function. The above equation is iteratively updated through gradient descent minimization as

$$C^t = -\delta(C) \omega_2 ((f(x) - \Phi^{(I)})^2 - (f(x) - \Phi^{(E)})^2) - \omega_1 \nabla \cdot \frac{\nabla C(x)}{|\nabla C(x)|} \quad (3)$$

Here $\delta(C) = \frac{1}{\pi} \cdot \frac{\varepsilon}{\varepsilon^2 + C^2}$ is the delta function and iterative adaptations of $\Phi^{(I)}$ and $\Phi^{(E)}$ are initiated with

$$\Phi^{(I)} = \frac{\iint f(x) M(C(x)) dx}{\iint M(C(x)) dx} \quad (4)$$

$$\Phi^{(E)} = \frac{\iint f(x) (1 - M(C(x))) dx}{\iint (1 - M(C(x))) dx} \quad (5)$$

Where $\Phi^{(I)}$ gives the intensity averages on the inside and $\Phi^{(E)}$ gives the same on the outside of the contour. The major drawback in CV model has been the assumption that the above intensities are homogeneous. This assumption fails to characterize image pixels globally. Hence, pixel intensity inhomogeneity segmentation problem persists in CV models which results in improper segmentations which has shown improvement using shape prior CV [39]. However, texture of the object in real time applications plays an important health detection factor and hence we propose our model called successive texture and shape-based AC model (STSAC).

B. STSAC Model

This section presents a new AC model that fits the evolving contour to the objects in an image based on additional information in the form of texture apart from previously used shape image. Consequently, a serial texture and shape influenced level set model is being formulated by describing each model separately.

1) *Texture Features*: Here, texture features are extracted using the most successful texture abstraction algorithm called Local Binary Patterns (LBP) [40]. The algorithm operates in a set pixel neighbourhood to identify the texture locally across the entire image. The central pixel in a region is compared with the neighbourhood pixels to binarize the region around the central pixel.

For a colour video frame $v(x, 3) \in R^+$, where x gives pixel location and the number 3 represents three colour planes RGB. The algorithm uses either 3 ($r=1$) or 12 ($r=2.5$) pixel neighbourhoods. The lbp texture code is modelled across a central pixel $(x_c, 3)$ as

$$lbp(x_c, 3) = \sum_{i=1}^3 \sum_{j=1}^P s(g_p - g_c) 2^P \quad (6)$$

Where,

$$s(x) = \begin{cases} 1 & \forall x \geq 0 \\ 0 & \text{Otherwise} \end{cases} \quad (7)$$

The g_c denotes the gray value at (x_c) and g_p represents the gray value in the neighbourhood of g_c . The variable P is the number of pixels around g_c . Using the above texture function, the active contour texture energy functional is formulated as

$$E_{texture}(\varphi) = d^2(\varphi_0, \varphi_t) = \iint_C (-H(\varphi_0) + H(F_{tf}))^2 \delta(x) dx \quad (8)$$

The term $F_{tf} = lbp(v(x))$.

2) *Shape Image*: To achieve clean object boundaries on bogie video data, in addition to texture, the chapter proposes to use shape knowledge. Here, the shape is modelled as a zero-active contour for the first video frame of the bogie video sequence. Signed distance function computes the relation between the shape prior and the initial contour. Eventually, the evolving contour computes this signed distance function to move towards the zero contour through gradient minimization.

The energy functional with prior shape knowledge is given as

$$E^S(C, \phi_S^{(I)}, \phi_S^{(E)}) = \iint_C (H(\phi(x)) - H(\phi_S(x)))^2 \delta(\phi) dx \quad (9)$$

Where E^S gives the shape energy function of the active contour. The energy function in eq'n (9) is minimized iteratively to find the object boundaries through an initial contour $f : c \rightarrow R^2$ to $\phi_S^{(I)}$ and $\phi_S^{(E)}$. Where $\phi_S^{(I)}$ and $\phi_S^{(E)}$ define the inside and outside portions to shape boundary.

3) *Level Set Formulation*: The intensity inhomogeneity in train bogie videos for accurate segmentation of bogie parts with both shape and texture measures simultaneously can be formulated as a serial texture shape based level set energy functional defined as

$$E^{S_TS} = E_{texture} + \tau \cdot E^{Shape} \quad (10)$$

Where τ is the delay between texture prior information and shape prior information applied as pre-knowledge to the ACs. The serial texture and shaped energy function is defined as

$$\begin{aligned} E^{S_TS}(C, \Phi^{(I)}, \Phi^{(E)}) = & \min_{C, \Phi^{(I)}, \Phi^{(E)}} \omega_2 \left[\iint_{C^I} (f(x) - \phi^{(I)})^2 H(C(x)) \right. \\ & \left. + \iint_{C^E} (f(x) - \phi^{(E)})^2 (1 - H(C(x))) dx \right] \\ & + \omega_1 \int_C |\nabla H(C(x))| dx \\ & + \lambda \iint_C (-H(\varphi_0) + H(F_{tf}))^2 \delta(x) dx \\ & + \tau \cdot \xi \iint_C (H(\phi(x)) - H(\phi_S(x)))^2 \delta(\phi) dx \end{aligned} \quad (11)$$

where λ and ξ are the controls for the texture and shape which can be applied as prior information. All the variables and parameters in eq (11) have the same representation as the previous model in Section III.A. The contour evolution is achieved by applying the gradient descent model to achieve a minimization as

$$\begin{aligned} C^t = & \left[\xi (g_{tex}(|\nabla f_x|) |\nabla \phi(x)|) \delta(\phi) \right. \\ & \left. + \gamma \left((f_x - \phi^I)^2 + (f_x - \phi^E)^2 \right) \delta(\phi) + \mu \nabla \cdot \frac{\nabla C(x)}{|\nabla C(x)|} \delta(\phi) \right] \\ & + \tau \cdot \lambda (H(\phi(x)) - H(\phi_S(x)))^2 \delta(\phi) \end{aligned} \quad (12)$$

Where g_{tex} is the texture function defined using local binary pattern (LBP). The proposed serial texture shape active contour model has shown improved performance in the segmentation of objects in real time video sequences that are inhomogeneous with the surroundings.

4) *STSAC Contour evolution*: For contour evolution in eq (12), is implemented on a machine using the model from [39], which is given by

$$C_{new} = C_{old} + \Delta t \frac{dC}{dt} + \Delta t C_{Texture} \quad (13)$$

After the texture evolution is completed when the following stopping criteria $\|\nabla \phi^n(x, y)\| \leq \varepsilon$ is attained. Where ε is the minimum gradient between the last two consecutive contour

evolutions in texture prior active contour model. Once, the texture prior stops, the outer edges formed due to texture evolution show a poor boundary or shape representation due to inhomogeneity in the object boundaries. To reconstruct the shape of the textured segmentation, we now apply the shape prior model on the textured contour. This is unlike the previous models, where the texture and shape information are fused into a single prior model for the active contour. This single prior model work well if the texture and shape models are perfectly aligned in 2D space. However, in real time computer vision applications getting a perfectly fused texture and shape model is quite a difficult process. Hence, our proposed serial texture shape based contour evolution can handle both texture and shape influence accurately than the previous models for real time computer vision applications. The shape contour evolution

$$C_{new} = \left(C_{Old} + \Delta t \frac{dC_{Old}}{dt} \right)^{Texture} + \Delta t C_{Shape} \quad (14)$$

Where $\frac{dC}{dt}$ predicts the rough variations in the right-hand side of the eq's(13, 14) and $\Delta t = \frac{0.48}{\max(C_{Old})}$ gives the step size in time. Fig.2 shows the comparison of STSAC against the previous texture shape fused models.

Fig. 2(g) and (h) are the outputs from the serial textured shape based active contour model (STSAC), which are better than the previously proposed models. In the next section we present the experiments and related analysis of the proposed method STSAC for TRSE video datasets in Table I.

TABLE I. VIDEO DATASETS CAPTURED FOR TESTING THE PROPOSED METHODS.

Experiments	Name	Number of Frames
D-1/E-1	Bogie Video Recorded at 6.40AM	90 × 36 = 3240
D-2/E-2	Bogie Video Recorded at 12.40PM	90 × 40 = 3600
D-3/E-3	Bogie Video Recorded at 4.20PM	90 × 32 = 2880
D-4/E-4	Bogie Video Recorded at 6.50PM	90 × 26 = 2340
D-5/E-5	Defective Video on 12.40PM Train	40 × 4 + 90 × 38 = 3600

IV. RESULTS AND DISCUSSION

This section discusses the datasets capturing mechanism and their characteristics in detail. Next, an extensive experimentation of the proposed algorithm on the considered TRSE problem for extracting bogie parts is presented. The results obtained are evaluated and analysed with benchmark algorithms already proposed on TRSE.

A. The Datasets

To test the proposed methods for their ability in extracting objects from videos captured under ambient conditions, the train rolling stock video database is built. The videos are recorded near to an Indian Railway station using the setup shown in Fig. 3. The figure shows an arrangement not more than 3 feet from the moving train. All the videos were recorded when the train was entering the station for a halt.

The handbook on train rolling stock examination was followed during the video capture. Accordingly, all trains in the dataset were recorded when the train was running at around 30KMPH. However, Digital single lens reflex (DSLR) record

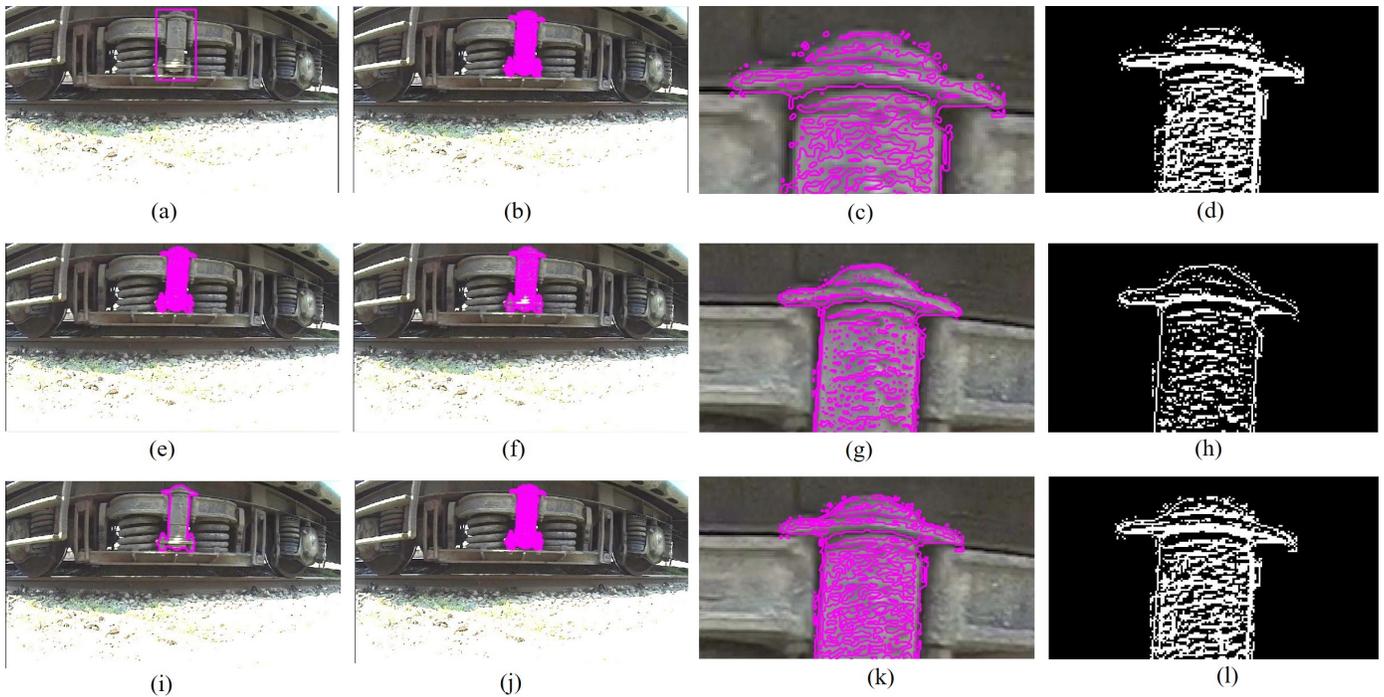


Fig. 2. Comparison of only texture and shape prior based active contour models with the proposed serial texture shape based active contour model. (a) Initial contour, (b) Only texture based AC model after 50 iterations, (c) Zoomed View of the object from (b), (d) Zoomed view of the segmented result from (b), (e) Our proposed initial texture model, (f) serial shape model, (g) Zoomed from (f), Segmented result from the proposed model STSAC in (f), (i) shape only AC model and (j) fused shape texture model, (k) zoomed shape textured model, (l) zoomed segmented output from (k).

at around 30 frames per second(fps), which induce a consider amount of motion blur to the video object data. Hence, to do away with blurring, the videos were recorded with Isaw sports action camera as can be seen in Fig. 3.



Fig. 3. Train Bogie Video Recording using High Speed Sports action Camera.

The Isaw sports action camera captures video at a max frame rate of 240fps. The camera also possesses a wide-angle lens with a 52° angle, that is capable of full bogie into a video frame from the centre. Fig. 4 shows an array of bogie video frames that were recorded by the visual sensor near the tracks. A total set of 4 train bogie videos were filmed at separate time stamps on a day. The advantage of in this approach gives an opportunity to test the proposed methods ability to overcome the effects of ambient lighting on segmentation quality.

In this paper, operation on each video sequence is considered as an experiment. Hence, 4 experiments were performed for testing the proposed methods. Moreover, a 5^{th} experiment is added to test the proposed algorithms capabilities in



Fig. 4. Video Frames of a recorded Bogie for Experimentation.

segmenting defective bogie parts. Unfortunately, there were no comprehensive defects in the recorded videos. Hence, the 5^{th} video has been handcrafted by extracting frames and deliberately inducing defects. Thereupon the defective frames were incorporated back to form a defective train sequence which resulted in a defective train bogie video for the 5^{th} experiment.

Fig. 5 shows the video frames that have been photoshopped with defects to bogie parts. The purpose of the proposed segmentation algorithms is resolved to satisfaction if it manages to segment the defective part through the prior knowledge of the healthy bogie part. This capability of the proposed frameworks in this work increases the scope for automation. Finally, Table I shows the experimental valuations performed on the five different datasets throughout the thesis. Fig. 6 gives a visualization of the datasets from Table I.

B. Bogie Parts Segmentation

Fig. 7 projects the results of the segmentation on bogie video frames for the 10 different parts as shown in column 1.



Fig. 5. Bogie part defects induced with photoshop. (a) Spring breaks and (b) Binding rod breaks.

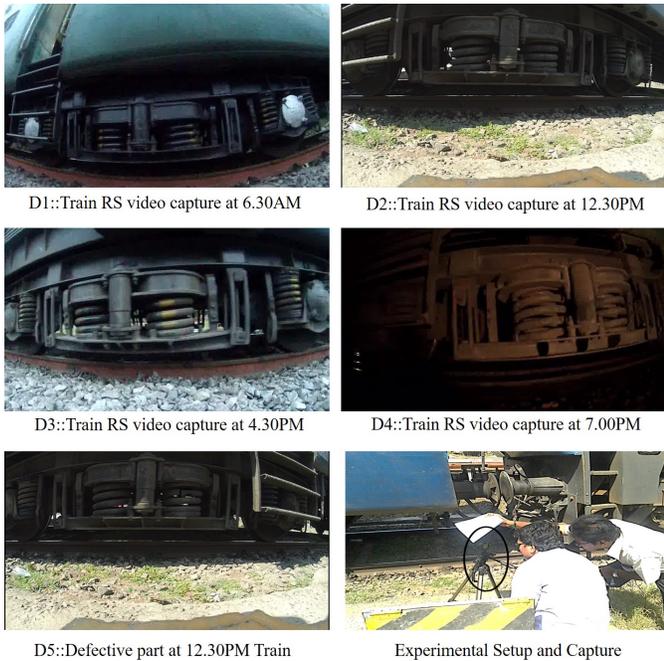


Fig. 6. Visualization of datasets.

Visual analysis of the Fig. 7 shows that the proposed STSAC segments bogie parts by retaining both shape and texture. The STSAC method requires the selection of bogie part that the user needs to check from the recorded video sequence along with the initial location of object. Fixing these hyperparameters can automate the part extraction process using STSAC.

The first column in Fig. 7 is the shape prior model extracted from the train bogie frame. Texture extracted using LBP is in second column and the next four columns show the curve evolution for all ten bogie parts after fixed number of 15 iterations. The last column shows the segmented bogie parts from the proposed STSAC algorithm. A visual comparison with similar algorithms is shown in Fig. 2.

The standard parameters ξ , γ and μ are kept constant and low for smooth contour propagation on the video frame. The initial texture contour evolution stopped in most of the cases at around 28th iteration and from then on, the shape evolution lasted for 56th iteration.

Compared to the shape or texture only based models the computation cost is little on the higher side for STSAC model as it had to run serially. The maximum number of iterations recorded were 68, in case of springs. For the entire train of 17 bogies in our dataset 1, our program took 0.35 hours for evaluation on an 8GB RAM with 2.4GHz intel processor on MATLAB software.

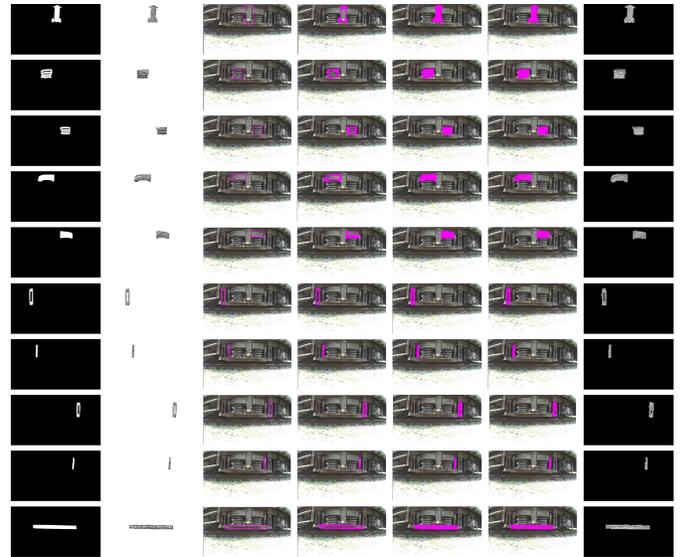


Fig. 7. Bogie part segmentation outputs on the 1st dataset using the proposed STSAC model. (Zoom In for better visibility).

The results demonstrate that the STSAC is capable of segmenting the region of interest objects given their texture and shape information accurately. However, from Fig. 1, we see that the similar AC models with only texture or shape or both had lost either texture or shape in the final output segment on real time video data. Our proposed model has retained good amount of texture information when shape is being reconstructed during the curve evolution process. The segmented outputs could be evaluated further for identifying their defects or their running life by comparing them with the available reference models.

This part of the work is achieved by comparing the segmented bogie part with the reference parts captured from the railway workshop. The reference parts are binarized and are called Ground Truth frames. These images appear as the first row in Fig. 7.

C. Defect Detection

Fig. 8 shows the defect in the bidding screw and the contour evolution using the proposed algorithm to identify the defect.

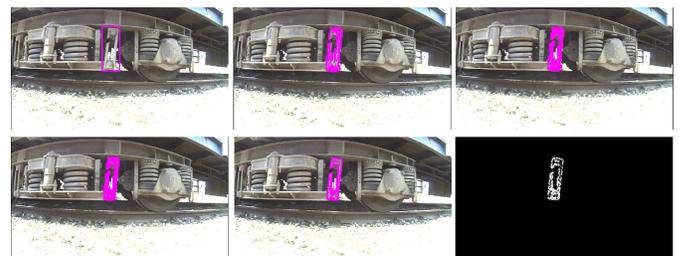


Fig. 8. Defect identification using the proposed STSAC model.

The previous models came close to the above result but showed multiple regions around the defect region making it difficult to identify the actual breakage point as shown

in fig.8. Similar results were obtained for cut defects of dimension 0.2mm thick on all bogie parts. For defects less than 0.2mm, our proposed model failed to extract the defect in the segmented output.

D. Bogie Video Data Analysis

This section parametrically evaluates the performance of the previously proposed algorithms against the proposed STSAC model across the datasets in Table I. The following parameters are computed for benchmarking the performance of the proposed algorithm against state-of-the-art previous models.

Detection sensitivity (DS) is defined as the parameter for measuring the performance of the segmentation methods. DS calculates the number of times an object is segmented correctly in the entire video sequence against the times it could not. The number of instances in the entire video sequence an object has been identified or segmented accurately is called True Positive (TP). Similarly, the opposite of TP is True Negative (TN), which gives the number of instances an object of interest could not be segmented properly. The expression for DS is

$$DS = \frac{TP}{TP + FN} \quad (15)$$

A DS measure of 0 indicates a failed segmentation and a one indicates a successful segmentation. The parameter is important for the proposed TRSE application to measure the accuracy of the methods in segmenting the required bogie parts in the entire video sequence.

The Mean Absolute Distance (MAD) calculates the deviation in the segmentation result of an algorithm from the actual required output. It is calculated by subtracting the obtained result with the ground truth GT of the object given as

$$MAD = \left| \frac{|S^O - GT| - \sum_{i=1}^n |S^O - GT|}{n} \right| \quad (16)$$

where 'n' is the number of pixels.

The Boundary Mean Absolute Distance (BMAD) parameter illustrates the deviation between boundaries of resulting segmented objects S_b^O to ground truth GT_b . The following expression for BMAD is

$$BMAD = 1 - \left| \frac{|S_b^O - GT_b| - \sum_{i=1}^n |S_b^O - GT_b|}{n_{Boundary}} \right| \quad (17)$$

Where, ' $n_{Boundary}$ ' is the number of pixels in the GT object boundary.

The Normalized Mutual Information (NMI) is a measure to determine how close the resulting segmented object is to the ground truth object. It gives a degree of common information in images. NMI is given as

$$NMI = \frac{MI(I, GT)}{\sqrt{H(I)H(GT)}} \quad (18)$$

Where $MI(I, GT) = - \sum_{I, GT} p(I_i, GT_i) \cdot \log \left(\frac{p(I_i, GT_i)}{p(I_i)p(GT_i)} \right)$ is the mutual information and $H(\bullet)$ is the entropy. MNI has a scale of 0 to 1, where 1 means highest segmentation accuracy.

Finally, the train rolling stock examination is being performed as a real time operation. Consequently, it becomes necessary for the proposed algorithms to compete in speed of execution. Hence, Number of Iterations – Model Speed (MS) measures the number of iterations in which the initial contour deforms and encompasses the object of interest.

Initially, first two parameters, DS and MAD are calculated and averaged on a set of 2000 bogie video frame segmentation outputs. Table II presents the computed values of various AC models on our rolling video datasets in Table I. The parameters are average across datasets. The values point to a conclusion that there has been a direct link between DS and the ambient lighting in which videos are recorded. Similarly, it can be seen that the STSAC performed quite well over the AC models with only shape (SP_AC), texture (TP_AC) and fused shape texture (FSTP_AC) prior models.

E. Parametric Analysis Against the State-of-the-Arts

Here, the proposed STSAC model is being validated against the state-of-the-art AC models parametrically with DS, MAD, BMAD, NMI and NI defined in the previous section. The computed parameters are plotted in Fig. 9. Fig. 9 plots the average DS of the localized AC models shape, texture and fused prior models and our proposed STSI_AC for all the bogie parts. The values plotted in Fig. 9 are averaged across all 4 datasets from Table I. The proposed approach showed that it can detect texture in inhomogeneous regions provided a weak shape and texture prior model as references.

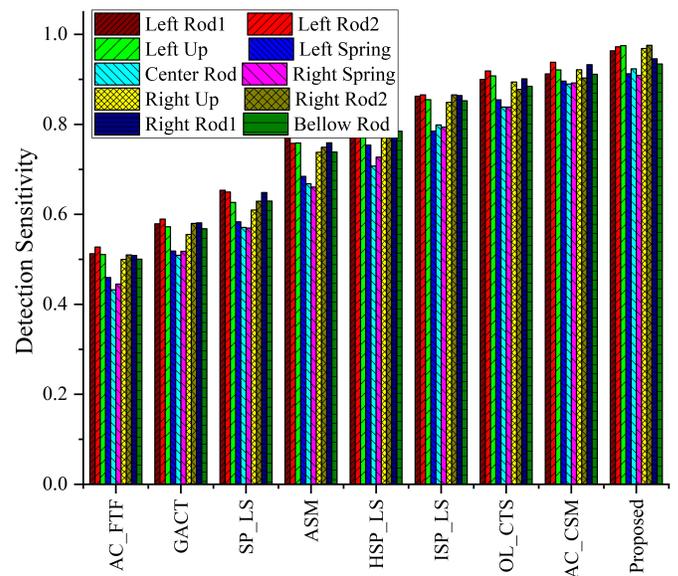


Fig. 9. Detection sensitivity of the proposed approach (STSAC) with the state of the art similar models.

The other plots for the remaining four parameters are presented in Fig. 10(a), (b), (c) and (d) show that the proposed STSI_AC is the better segmentation algorithm in the batch.

All the parameters are computed like the previous chapter. The proposed STSAC has been successfully implemented to extract texture and shape serially using AC models. However, the occlusion resistance has been on the lower side when compared to the shape prior models. Apart from that the model in this work provides additional information on the segmented bogie part for better decision making on the quality of the object.

segment a train bogie part for inspection using the high-speed video data of the train moving at 30KMPH. The video of the moving train was captured with a high – speed camera at 240fps. Serial texture shape prior active contours algorithm has been developed which uses first the texture prior and then the shape prior serially to extract objects texture by preserving its shape. This was quite different from similar algorithms which uses either texture or shape or both in fused form as priors, resulting in less than accurate segmentation outputs on real time video data. However, the proposed model had bettered the segmentation outputs both visually and parametrically over the existing models. Hence, the proposed method shows prospects of inducing as a platform for segmenting train bogie parts for automated train rolling stock examination in real time.

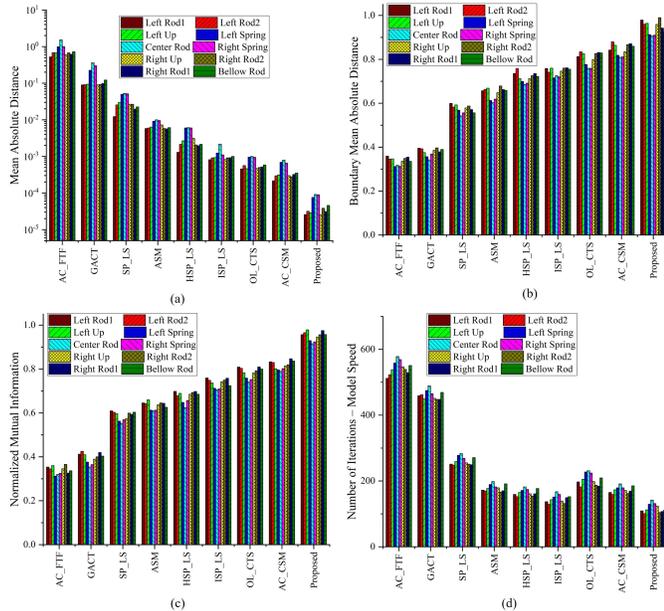


Fig. 10. Performance measures of the proposed algorithm against the state-of-the-art models for automated train rolling stock video segmentation: (a) Mean absolute distance, (b) Boundary mean absolute distance, (c) Normalized mutual information and (d) Number of iterations.

V. CONCLUSION

To improve the quality of the segmented bogie parts, this thesis proposes the second active contour model called serial texture and shape influenced active contour (STSAC). Traditionally, texture and shape based active contours use texture information in a region for shape segmentation. In contrast, this work proposes to segment shape rich texture objects through contour evolution performed serially on a shape prior model. This resulted in an improvement in segmented bogie parts over the previous model. This work presents a real-time computer vision problem and generates a formidable solution using novel serial texture shape prior active contour models. The objective of the real-time computer vision problem is to

REFERENCES

- [1] Maik Rosenberger and Rafael Celestre. Smart multispectral imager for industrial applications. In *2016 IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 7–12. IEEE, 2016.
- [2] Petra Tatzert, Markus Wolf, and Thomas Pannier. Industrial application for inline material sorting using hyperspectral imaging in the nir range. *Real-Time Imaging*, 11(2):99–107, 2005.
- [3] Frank Joachim Grote and Carsten Buchwald. Beverage bottling plant having an apparatus for inspecting bottles or similar containers with an optoelectric detection system and an optoelectric detection system, December 20 2016. US Patent 9,522,758.
- [4] Hong-Dar Lin and Kuan-Shen Hsieh. Automated distortion defect inspection of curved car mirrors using computer vision. In *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (ICCV)*, page 361. The Steering Committee of The World Congress in Computer Science, Computer ..., 2015.
- [5] H Golnabi and A Asadpour. Design and application of industrial machine vision systems. *Robotics and Computer-Integrated Manufacturing*, 23(6):630–637, 2007.
- [6] Desmond K Moru and Diego Borro. A machine vision algorithm for quality control inspection of gears. *The International Journal of Advanced Manufacturing Technology*, 106(1-2):105–123, 2020.
- [7] Jay H Heizer and Barry Render. *Operations management*, volume 1. Pearson Education India, 2008.
- [8] Fei Yan, Chunhai Gao, Tao Tang, and Yao Zhou. A safety management and signaling system integration method for communication-based train control system. *Urban Rail Transit*, 3(2):90–99, 2017.
- [9] Michal Kuffa, Daniel Ziegler, Thomas Peter, Fredy Kuster, and Konrad Wegener. A new grinding strategy to improve the acoustic properties of railway tracks. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 232(1):214–221, 2018.
- [10] Xiaoxuan Wang, Hailin Jiang, Wenzhe Sun, and Tao Tang. Efficient dual-association resource allocation model of train-ground communication system based on td-lte in urban rail transit. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 2006–2011. IEEE, 2016.

TABLE II. AVERAGE DS AND MAD FOR EACH DATASET WITH THE PROPOSED MODEL AND THE STATE OF THE ART LEVEL SET MODELS.

Methods	D-1		D-2		D-3		D-4		D-5	
	DS	MAD								
AC_FTF [37]	0.45	0.89	0.49	0.86	0.51	0.84	0.32	0.95	0.39	0.82
GACT [41]	0.44	0.91	0.51	0.81	0.52	0.79	0.35	0.93	0.33	0.84
SP_LS [28]	0.58	0.84	0.59	0.8	0.59	0.81	0.4	0.9	0.44	0.89
ASM [42]	0.79	0.39	0.78	0.32	0.77	0.3	0.72	0.35	0.7	0.3
HSP_LS [43]	0.68	0.47	0.71	0.52	0.75	0.56	0.42	0.55	0.43	0.86
ISP_LS [27]	0.7	0.43	0.75	0.44	0.78	0.42	0.51	0.48	0.45	0.87
OL_CTS [44]	0.75	0.41	0.75	0.4	0.75	0.4	0.43	0.31	0.47	0.82
AC_CSM [45]	0.84	0.34	0.88	0.38	0.89	0.33	0.47	0.32	0.82	0.31
STSAC (Proposed)	0.91	0.31	0.94	0.25	0.92	0.28	0.83	0.34	0.88	0.35

- [11] Yong Qin, Baojun Yuan, and Si Pi. Research on framework and key technologies of urban rail intelligent transportation system. In *Proceedings of the 2015 International Conference on Electrical and Information Technologies for Rail Transportation*, pages 729–736. Springer, 2016.
- [12] Arzoo Naghiyev, Sarah Sharples, Brendan Ryan, Anthony Coplestone, and Mike Carey. Expert knowledge elicitation to generate human factors guidance for future european rail traffic management system (ertms) train driving models. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 231(10):1141–1149, 2017.
- [13] Satish Chandra et al. *Railway engineering*. Oxford University Press, Inc., 2008.
- [14] JS Mundrey. *Railway track engineering*. Tata McGraw-Hill Education, 2009.
- [15] Ashutosh Kumar Banerji. Railways and rolling stock engineers-challenges ahead. *Technical Note*, 2005.
- [16] Konkan railway automated train examination system, 2008. <https://www.youtube.com/watch?v=8RU54XZc9so>.
- [17] Sundaravalli Narayanaswami. Urban transportation: innovations in infrastructure planning and development. *The International Journal of Logistics Management*, 2017.
- [18] Alessandro Sabato and Christopher Niezrecki. Feasibility of digital image correlation for railroad tie inspection and ballast support assessment. *Measurement*, 103:93–105, 2017.
- [19] John M Hart, Esther Resendiz, Benjamin Freid, Steven Sawadisavi, CPL Barkan, and N Ahuja. Machine vision using multi-spectral imaging for undercarriage inspection of railroad equipment. In *Proceedings of the 8th world congress on railway research, Seoul, Korea*, volume 18, 2008.
- [20] HyunCheol Kim and Whoi-Yul Kim. Automated inspection system for rolling stock brake shoes. *IEEE transactions on instrumentation and measurement*, 60(8):2835–2847, 2011.
- [21] Angel Luis Sanchez-Revuelta and Carlos-Javier Gomez Gomez. Installation and process for measuring rolling parameters by means of artificial vision on wheels of railway vehicles, September 15 1998. US Patent 5,808,906.
- [22] NL Kazanskiy and SB Popov. Integrated design technology for computer vision systems in railway transportation. *Pattern Recognition and Image Analysis*, 25(2):215–219, 2015.
- [23] Benjamin Freid, Christopher PL Barkan, Narendra Ahuja, John M Hart, Sinisa Todorovic, and Nicholas Kocher. Multispectral machine vision for improved undercarriage inspection of railroad rolling stock. In *Proceedings of the Ninth International Heavy Haul Conference Specialist Technical Session—High Tech in Heavy Haul, Kiruna, Sweden*, pages 11–13, 2007.
- [24] Leszek Jarzebowski and Slawomir Judek. 3d machine vision system for inspection of contact strips in railway vehicle current collectors. In *2014 International Conference on Applied Electronics*, pages 139–144. IEEE, 2014.
- [25] Yu Zhang, Jia-yuan Hu, Jin-long Li, Hai-qing Wang, et al. The application of wtp in 3-d reconstruction of train wheel surface and tread defect. *Optik*, 131:749–753, 2017.
- [26] PVV Kishore and Ch Raghava Prasad. Computer vision based train rolling stock examination. *Optik*, 132:427–444, 2017.
- [27] Ch Raghava Prasad and PVV Kishore. Performance of active contour models in train rolling stock part segmentation on high-speed video data. *Cogent engineering*, 4(1):1279367, 2017.
- [28] PVV Kishore and Ch Raghava Prasad. Shape prior active contours for computerized vision based train rolling stock parts segmentation. *International Review on Computers and Software (I. RE. CO. S.)*, 10:1233–1243, 2015.
- [29] PVV Kishore and Ch Raghava Prasad. Train rolling stock segmentation with morphological differential gradient active contours. In *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1174–1178. IEEE, 2015.
- [30] Farhan Riaz, Sidra Naeem, Raheel Nawaz, and Miguel Coimbra. Active contours based segmentation and lesion periphery analysis for characterization of skin lesions in dermoscopy images. *IEEE journal of biomedical and health informatics*, 23(2):489–500, 2018.
- [31] Tiejun Yang, Yaowen Chen, and Zhun Fan. Vegetation segmentation based on variational level set using multi-channel local wavelet texture and color. *Signal, Image and Video Processing*, 12(5):951–958, 2018.
- [32] Shiyu Luo, Ling Tong, and Yan Chen. A multi-region segmentation method for sar images based on the multi-texture model with level sets. *IEEE Transactions on Image Processing*, 27(5):2560–2574, 2018.
- [33] Nawal Houhou, Jean-Philippe Thiran, and Xavier Bresson. Fast texture segmentation model based on the shape operator and active contour. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008.
- [34] Mingqi Gao, Hengxin Chen, Shenhai Zheng, and Bin Fang. Feature fusion and non-negative matrix factorization based active contours for texture segmentation. *Signal Processing*, 159:104–118, 2019.
- [35] Priyambada Subudhi and Susanta Mukhopadhyay. A novel texture segmentation method based on co-occurrence energy-driven parametric active contour model. *Signal, Image and Video Processing*, 12(4):669–676, 2018.
- [36] Guo Zhao, Shiyin Qin, and Danyang Wang. Interactive segmentation of texture image based on active contour model with local inverse difference moment feature. *Multimedia Tools and Applications*, 77(18):24537–24564, 2018.
- [37] Qinggang Wu, Yong Gan, Bin Lin, Qiuwen Zhang, and Huawen Chang. An active contour model based on fused texture features for image segmentation. *Neurocomputing*, 151:1133–1141, 2015.
- [38] Tony F Chan and Luminita A Vese. Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277, 2001.
- [39] Tony Chan and Wei Zhu. Level set based shape prior segmentation. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 1164–1170. IEEE, 2005.
- [40] Syed Inthiyaz, BTP Madhav, and PVV Kishore. Flower segmentation with level sets evolution controlled by colour, texture and shape features. *Cogent Engineering*, 4(1):1323572, 2017.
- [41] Berta Sandberg, Tony Chan, and Luminita Vese. A level-set and gabor-based active contour algorithm for segmenting textured images. In *UCLA Department of Mathematics CAM report*. Citeseer, 2002.
- [42] Mina Esfandiarkhani and Amir Hossein Foruzan. A generalized active shape model for segmentation of liver in low-contrast ct volumes. *Computers in biology and medicine*, 82:59–70, 2017.
- [43] Xiong Yang, Shu Zhan, Dongdong Xie, Hong Zhao, and Toru Kurihara. Hierarchical prostate mri segmentation via level set clustering with shape prior. *Neurocomputing*, 257:154–163, 2017.
- [44] Yu Zhong and Anil K Jain. Object localization using color, texture and shape. *Pattern Recognition*, 33(4):671–684, 2000.
- [45] Weiming Hu, Xue Zhou, Wei Li, Wenhan Luo, Xiaojin Zhang, and Stephen Maybank. Active contour-based visual tracking by integrating colors, shapes, and motions. *IEEE Transactions on Image Processing*, 22(5):1778–1792, 2012.