Invariant Features-Based Fuzzy Inference System for Animal Detection and Recognition Using Thermal Images

S. Divya Meena[1] ⬤ · **L. Agilandeeswari**[1]

Abstract Human–Animal Conflict (HAC) is one of the primary threats to the continued survival of animal species and it has also impacted the lives of humans drastically. In this paper, we propose an efficient animal detection and recognition system with invariant features and fuzzy logic using thermal images. The proposed system exploits various features like Zernike, shape, texture and skeleton path. Cumulatively, these features are invariant to rotation, scaling, translation, illumination, and partly posture. The proposed model is robust to several challenging image conditions like low contrast/illumination, haze/blur, occlusion, camouflage, background clutter, and poses variation. The model is tested on our thermal animal dataset that has 1862 images and 12 different animal species. Experimental results validate the significance of thermal images for animal-based applications. Besides, the proposed fuzzy system has achieved an average accuracy of 97% which is equivalent to the accuracy produced by domain experts in identifying the animals from our thermal dataset.

Keywords Fuzzy logic · Thermal images · Skeleton and shape features · Zernike moments

1 Introduction

Wild animal monitoring systems are gaining importance due to the number of Human–Animal Conflict (HAC) that has been occurring over the decades. With visible images, it is tedious to detect animals during night time due to the engulfed darkness. Besides, animals could self-mask with their flexible structure and their cluttered background adds to the complexity. However, thermal imaging cameras are one of the perfect tools for night vision applications. They work on the principle of heat energy and so they can detect hot-blooded organisms like human, animals very easily, besides ignoring camouflage. Although camera traps are one of the best tools for capturing the animals, they do not always capture perfect images and they do pose several challenges like, noise, low contrast/illumination, haze/blur, occlusion, camouflage, background clutter, and pose variations leading to poor interpretation. As a counter measure, we extract the invariant features which are not deterred by these challenges. Invariant features are a special type of features which can identify the animals precisely even when they appear slightly different from their original form, thus making them robust to any challenging image conditions.

The number of animal detection models with thermal images is relatively low and the primary reason is the lack of publicly available thermal animal dataset. Yet, thermal images have been used in few animal studies like studying the population of Brazilian free-tailed bats [1, 2], detecting diurnal terrestrial reptiles [3], identifying large mammals in African bushveldt [4], welfare monitoring system of rodents [5] and Deer–Vehicle Collision detection system [6, 7]. Besides, Cilulko et al. [8] reviewed the various uses of thermal images in wildlife studies. Real-time vehicular animal detection system with thermal cameras was

✉ L. Agilandeeswari
  agila.l@vit.ac.in

[1] School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India

proposed for Autoliv cars [9] and is now used in Audi, BMW and Daimler cars. With the Fuzzy Inference System (FIS), we can distinguish the severity of a tiger cub from an adult tiger and this kind of crisp inference is highly essential in real-time applications involving human lives. Although fuzzy logic has been used in a wide variety of applications like text recognition [10], moving object detection [11], human detection [12], diagnosing disease in rice [13], detecting estrus in cows [14], pedestrian detection [15, 16], and vehicle detection [17]; it has not been studied on animal-based applications due to the complex characteristics of animals. This motivated us to propose the first of its kind fuzzy logic-based animal detection and recognition system using thermal images. The contributions of this work include.

(1) Developing an efficient animal detection and recognition system with invariant features and fuzzy logic using thermal image. The model is robust to several challenging image conditions.

(2) To investigate the influence of discrete features and feature fusion in detecting animals. Also, to investigate the effects of distance between the thermal camera and animal.

(3) To investigate other notable factors like weather conditions required for achieving best results with thermal images.

The remaining of the paper is organized as follows: Sect. 2 describes the proposed methodology. Section 3 discusses the experimental framework including results and discussion. Finally, the conclusion and future scope are presented in Sect. 4.

2 Proposed Methodology

In this section, we discuss image pre-processing followed by feature extraction and finally the fuzzy inference system for recognizing the animals from thermal images. The proposed model is entitled "*IF-FUZ*" with reference to the invariant features and fuzzy logic. Figure 1 depicts the flow of the proposed model.

2.1 Image Pre-processing

The raw images captured from the thermal camera are not appropriate for recognition. We follow a series of steps for pre-processing the image [18]. First, we use Gamma correction to correct the luminance of the image followed by histogram equalization to enhance the contrast. Next, we apply gradient-based guided edge-aware smoothing filter to smoothen the image. The edge-aware filter preserves the edges of the image and retains the structure of the animal
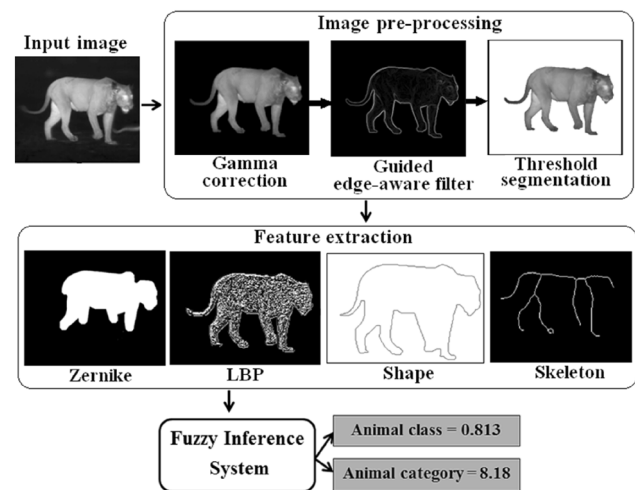

Fig. 1 Flow of proposed IF-FUZ

despite smoothing. Finally, we segment the object of interest by applying basic thresholding function. The outcome from each of the pre-processing step is given in the top portion of Fig. 1.

2.2 Feature Extraction

The feature extraction involves extracting the most informative data from the image. Different features have different capabilities and we consider a fusion of features. The first set of features to be extracted is the Zernike feature which is invariant to scaling, translation, and rotation. Animal do not pose for cameras and they can be in any form when captured. Hence, the invariant property of Zernike is the one of the most appropriate features for animals. The next feature captures the skeleton properties of the animal. In some cases, when the animals are deformed or occluded, the skeleton information is most informative as it is robust to shape noise, deformation and non-rigid transformations. Illumination problem is yet another factor affecting the animal recognition. In such cases, illumination invariant features like Local Binary Pattern (LBP) can be useful. Finally, we also extract the shape of the animal using a fusion of Statistical model-based methods (SMM) and Multi-index Active Model (MAM).

2.2.1 Zernike Features

Zernike moments which are based on the Zernike polynomial are orthogonal in nature and so invariant to rotation. The Zernike moment for a given ordered pair (m, n) is defined as
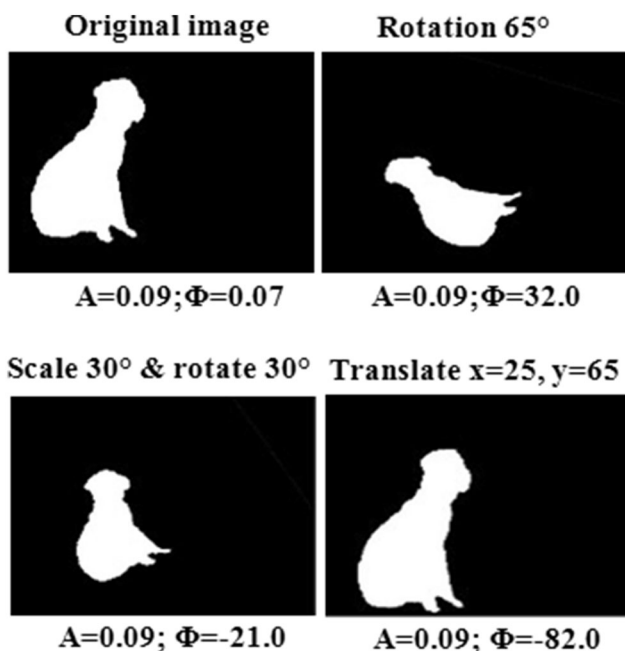
Original image

Rotation 65°

A=0.09;Φ=0.07

A=0.09;Φ=32.0

Scale 30° & rotate 30°

Translate x=25, y=65

A=0.09; Φ=-21.0

A=0.09; Φ=-82.0

Fig. 2 Invariance property of Zernike features

$$V_{nm}(\rho, \theta) = R_{nm}(\rho)e^{im\theta}, \theta \leq 1 \qquad (1)$$

where 'm' denotes the order of Zernike polynomial and 'n' the multiplicity/repetition of the phase angle. The ρ and θ in the above equation are given by

$$\rho = \sqrt{x^2 + y^2}; \theta = \arctan\left(\frac{y}{x}\right) \qquad (2)$$

where ρ is the radial vector of the image pixel and θ the angle. R_{nm} in Zernike moments is the Zernike polynomial and is given by

$$R_{nm}(\rho) = \sum_{a=0}^{\frac{(n-|m|)}{2}} (-1)^a \frac{(n-a)!}{a!\left(\frac{n+|m|}{2} - a\right)!\left(\frac{n-|m|}{2} - a\right)!} \rho^{n-2a}. \qquad (3)$$

For a given continuous function $f(x,y)$, the Zernike moment for n, m is given by

$$A_{nm} = \frac{n+1}{\pi} \sum_{\rho=0}^{1} R_{nm}(\rho) \sum_{\theta=0}^{2\pi} f(\rho.\theta)e^{-im\theta}. \qquad (4)$$

To have a complete feature set, we divide the image horizontally into two sub-regions (left and right) and combine the features from left $L(x_n, y_n)$ and right $R(x_n, y_n)$ into complete feature vector $I(x, y)$. We calculate the higher and lower order Zernike moments (see Table 1) using Eqs. 5 and 6.

$$\text{Lower order ZM} = \{Z_{n,m}\} \forall \begin{cases} 2 \leq n \leq 7 \\ |m| \leq n \\ n - |m| = 2k \\ k \in N. \end{cases} \qquad (5)$$

$$\text{Higher order ZM} = \{Z_{n,m}\} \forall \begin{cases} 7 \leq n \leq 12 \\ |m| \leq n \\ n - |m| = 4k \\ k \in N. \end{cases} \qquad (6)$$

2.2.2 Skeleton Feature

The skeleton structure is an object with many disks where only one disk will have maximal disk space and it is the locus of Center of Maximal Disks (CMD). We represent the skeleton graph in line form (see Fig. 3a) and the points on the skeleton are used to measure the distance and to check if the line segments are CMD or not. The skeletal
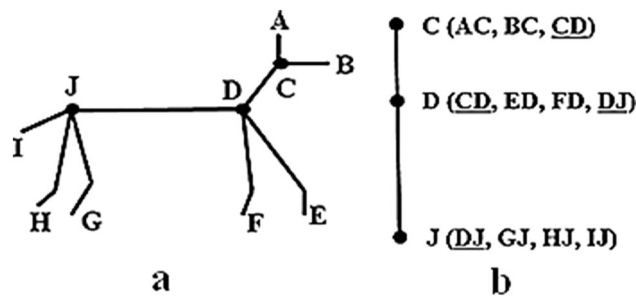
C (AC, BC, <u>CD</u>)

D (<u>CD</u>, ED, FD, <u>DJ</u>)

J (<u>DJ</u>, GJ, HJ, IJ)

Fig. 3 a Skeleton graph, **b** skeleton edge

Table 1 Higher and lower order Zernike moments

Group	N	M	Group	N	M
Lower order Zernike moments	2	0,2	Higher order Zernike moments	7	3,7
	3	1,3		8	0,4,8
	4	0,2,4		9	1,5,9
	5	1,3,5		10	2,6,10
	6	0,2,4,6		11	3, 7, 11
	7	1,3,5,7		12	0,4,8,12

Note that the amplitude of the Zernike moment (A) remains almost same despite changes in Fig. 2

point connected to only one other skeletal point is Ending node (E_{node}) and the skeletal point connected to three other skeletal points is Bifurcation node (B_{node}). In Fig. 2a, nodes A, B are E_{node} and nodes C, D, J are B_{node} and the remaining are normal nodes. Skeleton edges (SE) are formed by connecting the skeletal points between two B_{nodes} or two E_{nodes}. Primary SE is one with two B_{node} and all other nodes are normal edges. Node CD is primary SE and node AC is normal SE (see Fig. 3b).

The skeleton processing involves carefully extracting the skeleton as it is prone to shape noise and deformation. We prune unnecessary SE that occurs due to shape noise. To check if an SE is important or not, we compare each of the normal SE with original skeleton shape. If the difference is large, the SE is due to shape noise and it can be pruned; otherwise, it can be considered for the skeleton graph. The obtained skeleton can be further processed by representing the skeleton in tree form. The skeletal tree can be generated by tracing either the B_{node} or E_{node}. The nodes and edges for the tree are B_{node} and primary SE, respectively. The topology of 2D skeleton differs even for the same shape objects. To overcome this challenge, we use the skeleton path which is robust to shape noise. Let S be the skeleton and $e(S) = \{e_i, i = 1, 2, \ldots M\}$ be the end points of S where M denotes the total number of end points. In Fig. 3b, we have 10 end points. Skeleton path $p_{i,j}$ is the shortest path between the end points e_i and e_j and this is given by $p_{i,j} = (e_i, e_j)$. They are very informative and robust as they avoid unnecessary B_{node} through pruning. Moreover, despite the deformations and non-rigid transformations, skeleton path is same for similar shaped objects, as they depend on CMD rather than the topology of the object.

2.2.3 Local Binary Pattern (LBP)

The grayscale variant of thermal images may suffer from monotonic gray level changes called illumination variation. LBP is a highly discriminative texture descriptor that is invariant to illumination effects as illustrated in Fig. 4.

Fig. 4 Illumination invariance of LBP

LBP works by assigning a binary number to every pixels of the image. This binary number is typically called as label and it represents the relation between the center pixel and the corresponding neighboring pixel. This can be mathematically represented as

$$\text{LBP}(x_c, y_c) = \sum_{p=0}^{p-1} 2^p \times S(i_n - i_c) \tag{7}$$

where (x_c, y_c) is the center pixel, i_c its intensity and i_n the intensity of the neighboring pixel. S is the sign function which is equal to 1 if $x \geq 0$ else it is 0. The texture of the image can then be analyzed from the histogram of the labels.

2.2.4 Shape Feature Extraction

The animals have pose variation due to scaling, orientation, and rotation. To have a good recognition, the extracted features should be invariant to all the above. We utilize a hybrid variant shape extraction techniques based on SMM and MAM [19]. This method transforms the edge contour of the image to a specified position, orientation and scale. The variant methods are normally based on the reference shape, which is obtained by taking the average shape of few defectless images.

2.2.4.1 SMM-Based Shape Feature Extraction SMM relies on reference image for the shape feature extraction. The reference image is normalized by adjusting the position of the object to the center and also by adjusting the scale and orientation pertaining to the corresponding standard deviations. The intersection of x and y co-ordinates is the center point and the corresponding edge points are taken as the object's center. By sloping the major axis horizontally, we change the orientation. The original dimension (x_i, x_j) is scaled to new dimension (x_i', y_i') based on the standard deviation $(\text{SD}_x, \text{SD}_y)$ of the reference shape. New dimension is calculated by dividing the standard deviation from the old dimension. Further, we normalize all the images before comparing it with the normalized reference shape.

2.2.4.2 MAM-Based Shape Feature Extraction MAM feature extractor is also based on reference shape but it is an active method, as we can adjust the position, orientation and scale of the object in a way that it approximates the reference shape of the object. Shape transformation is illustrated with lion in Fig. 5. The outer layer is the actual edge contour of the original image. The inner most layer is the normalized edge contour and the shape near the normalized edge contour is the reference shape for the object. The line passing through the geometrical center, origin, tail
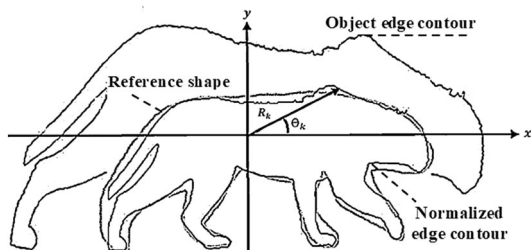
Fig. 5 Illustration of shape transformation

and the forehead of the lion is taken as *x*-axis. From the forehead of the line, a random number of equiangular positions are taken. The figure illustrates an arbitrary position (θ_k) and its corresponding radius (R_k). To compare the reference shape and the normalized object contour, we use several shape indices that are defined based on the relative difference of the radius, continuity and curvature. Radius index is the distance between the origin and the edge of the object contour and is defined as $I_{1,k} = R_k$. Continuity index is the measure of difference between the radiuses of any two adjacent equiangular positions and is given by $I_{2,k} = R_{k+1} - R_k$. Curvature index is a second-order derivative that follows the finite difference of the object contour at a given equiangular location and is defined as $I_{3,k} = R_{k-1} - 2R_k + R_{k+1}$. Area index is the surface area of the projected object and is denoted as I_4. Finally, aspect ratio is either the maximum breadth or length and is denoted as I_5. These five shape indices are calculated for both the reference shape $IR_{i,k}$ and the object under consideration $I_{i,k}$.

Animals have inherent variations along each of the equiangular position. To invalidate this, we consider the difference between $(I_{i,k})$ and $(IR_{i,k})$ with respect to the standard deviation of the reference shape $(SDR_{i,k})$ and it is defined as

$$D_{i,k} = \begin{cases} \dfrac{|I_{i,k} - IR_{i,k}|}{SDR_{i,k}}; & D_{i,k} > T \\ 0; & D_{i,k} \leq T \end{cases} \tag{8}$$

where *i* denotes the *i*th shape feature, *k* the *k*th equiangular position and *T* the threshold. In the experiments, we set *T*

as 1, so as to have a forbearance of 1 standard deviation around the reference shape contour. This has excluded minor object contour variations that arise due to inhomogeneity.

2.3 Fuzzy Inference System (FIS) for Animal Detection and Recognition

Fuzzy logic provides a flexibility with which human reasoning capability can be effectually applied to an artificial knowledge system. FIS has four major components namely fuzzification, rule base and defuzzification.

2.3.1 Fuzzification

In this step, we chose the inputs and their appropriate membership functions (MF), which is a mapping between the input (x) and the membership value $f(x)$. IF-FUZ have five inputs namely higher and lower order Zernike features, length of skeleton path, texture, and shape feature. The higher order Zernike feature is more accurate than lower order; however, it is prone to noise. Hence, we chose an optimal value for both higher (12th) and lower (4th) order. Of these 5 inputs, skeleton path and shape feature have triangular MF and the remaining has trapezoidal MF. Both the output variables utilizes the Gaussian MF, since it is non-zero and symmetric in all the points. We have defined two output variables namely the animal class and its category. Figures 6 and 7 depict the MFs of input and output variables, respectively.

Each of the inputs has three fuzzy sets. The first output variable (animal class) has 12 fuzzy sets representing the different animals and the second variable (category) has three fuzzy sets namely domestic, wild and others.

2.3.2 Fuzzy Rule Base

Fuzzy rule base is a set of verbal description based on the "if… then" rules. The fuzzy system arrives at a conclusion based on the rules we have defined. IF-FUZ has 48 rules as listed in Table 2. The rules were formulated based on the
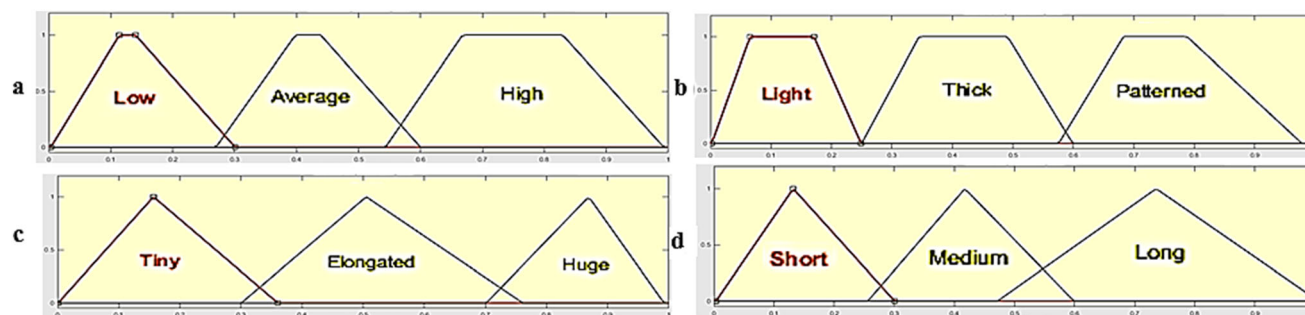

Fig. 6 Input membership function. **a** Zernike higher order, **b** texture, **c** skeleton path, **d** shape feature
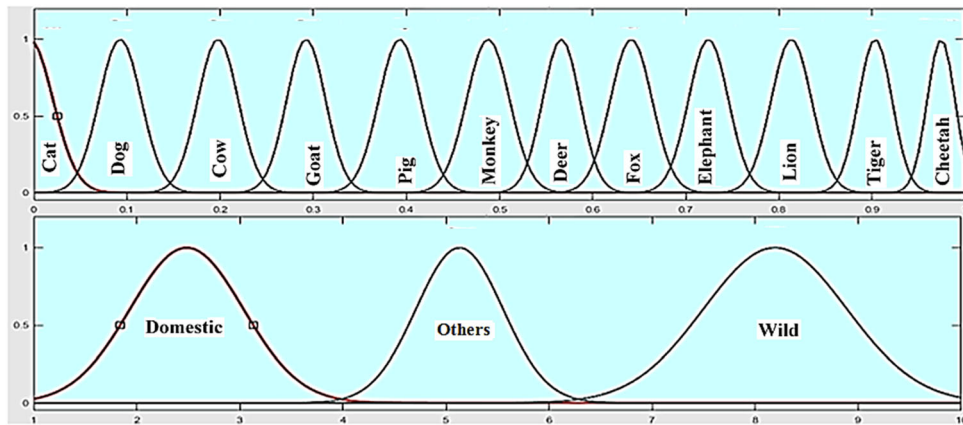
Fig. 7 Output membership function of animal class and their category

values obtained from the feature extraction techniques. The features were extracted and fixed on a range after normalizing the values. Upon few error and trials, the range for each of the features is defined properly.

2.3.3 Defuzzification

Defuzzification essentially quantifies the fuzzy information into numerical value. We finalized on using Centroid of Defuzzification (COD) after experimenting with various other techniques. Sample result is depicted in Fig. 8.

3 Experimental Results and Discussion

A part of the study was conducted in Vellore (12.9165° N, 79.1325° E), Tamil Nadu. The photos were taken all through the day to analyze the effects of thermal images during various times of the day. Few animals like fox, monkey and deer were captured from Amirthi Zoological Park, Vellore. The dataset with 1862 thermal images has 12 different animal classes that are clubbed into three different categories namely domestic, wild, and others. Under the wild category, animals such as lion, tiger, cheetah, and elephant were taken from the web. Except for these 4 classes, all other animals were manually captured with thermal camera. We used Forward Looking InfraRed (FLIR) thermal camera of model e40 [20]. FLIR e40 that works on Long-Wave Infrared Band (LWIR) has an IR resolution of 160 × 120 with a thermal sensitivity of 0.07° and temperature range of − 20° to 650°.

Sample images from the dataset and the challenging image conditions from camera traps are shown in Figs. 9 and 10, respectively.

3.1 Baseline Results

For the 1862 thermal images, a total of 46,550 detection and recognition attempts were made; where 25 volunteers assessed the 1862 images each. The 25 volunteers included 5 domain experts, 10 experienced volunteers and 10 novices. The detection attempts of domain experts were set as gold standard data and were used to cross-verify the results of all other volunteers. In sum, 66% of the attempts were correct; however, there was a substantial difference between various animal species. In case of elephant, 94% of attempts were correct, whereas only 37% of attempts were correct in the case of monkeys.

3.2 Performance Analysis of IF-FUZ

IF-FUZ is meticulously assessed with few standard performance measures like accuracy, precision and recall rate [21]. As the dataset is unbalanced, we also consider the F-measure, which is a weighted average value of Precision and Recall. The performance analysis of IF-FUZ on our thermal dataset is presented in Fig. 11.

IF-FUZ achieved an average accuracy of 97%, which is equivalent to the gold standard data produced by domain experts. In terms of precision, IF-FUZ had a positive predictive value of 91.4, which implies the model has been successful in recognizing the animals in most of the cases. The recall rate of the model is around 98%. We had 98.4% of weighted average value, thus signifying a good precision and recall rate.

IF-FUZ has used several feature descriptors and each of them has certain invariance property. To understand the importance of each features, we first tested the individual feature descriptors on the fuzzy model and then a fusion of them was tested. The results for the same are presented in Table 3. It could be inferred from the results that LBP and skeleton features are not sufficient for recognizing animals,

Table 2 Fuzzy rules for animal detection and recognition

| S. no | Input variable | | | | | Output variable | |
| | Trapezoidal MF | | | Triangular MF | | Gaussian MF | |
	Zernike_Low	Zernike_High	Texture	Skel_Path	Shape	Animal	Category
1	Low	Low	Light	Short	Tiny	Cat	Domestic
2	Low	Average	Light	Medium	Tiny	Cat	Domestic
3	Average	Low	Light	Short	Tiny	Dog	Domestic
4	Low	Average	Light	Medium	Elongated	Dog	Domestic
5	Average	High	Light	Medium	Elongated	Cow	Domestic
6	Average	Average	Light	Long	Huge	Cow	Domestic
7	Average	Low	Thick	Short	Tiny	Pig	Others
8	Low	Average	Thick	Medium	Tiny	Pig	Others
9	Average	High	Thick	Short	Tiny	Fox	Wild
10	Average	Low	Thick	Medium	Elongated	Fox	Wild
11	Average	Average	Patterned	Short	Tiny	Deer	Others
12	High	Average	Patterned	Medium	Elongated	Deer	Others
13	High	Low	Thick	Short	Tiny	Monkey	Others
14	Average	Average	Thick	Medium	Tiny	Monkey	Others
15	Low	Average	Light	Short	Tiny	Goat	Domestic
16	Average	Average	Light	Medium	Elongated	Goat	Domestic
17	High	Average	Thick	Long	Huge	Elephant	Wild
18	High	High	Thick	Long	Huge	Elephant	Wild
19	High	Average	Thick	Average	Elongated	Lion	Wild
20	High	High	Thick	Medium	Elongated	Lion	Wild
21	Average	Average	Patterned	Short	Elongated	Tiger	Wild
22	Average	Average	Patterned	Medium	Elongated	Tiger	Wild
23	High	Low	Patterned	Short	Elongated	Cheetah	Wild
24	Average	Average	Patterned	Medium	Elongated	Cheetah	Wild
25	Average	Low	Light	Short	Tiny	Cat	Domestic
26	Average	Average	Patterned	Medium	Tiny	Cat	Domestic
27	Average	Average	Light	Short	Tiny	Dog	Domestic
28	Average	High	Light	Medium	Elongated	Dog	Domestic
29	High	Average	Light	Medium	Average	Cow	Domestic
30	High	High	Light	Long	Elongated	Cow	Domestic
31	Average	Average	Light	Short	Tiny	Pig	Others
32	Average	High	Thick	Medium	Tiny	Pig	Others
33	Average	Average	Light	Medium	Tiny	Fox	Wild
34	High	Average	Thick	Medium	Elongated	Fox	Wild
35	Average	High	Thick	Medium	Tiny	Deer	Others
36	High	High	Light	Medium	Elongated	Deer	Others
37	Average	High	Thick	Short	Tiny	Monkey	Others
38	High	Low	Thick	Medium	Tiny	Monkey	Others
39	Average	High	Light	Short	Tiny	Goat	Others
40	High	Low	Light	Medium	Elongated	Goat	Domestic
41	Average	High	Thick	Long	Huge	Elephant	Wild
42	High	Low	Thick	Long	Huge	Elephant	Wild
43	Average	Low	Light	Short	Tiny	Lion	Wild
44	High	Average	Thick	Medium	Elongated	Lion	Wild
45	High	Average	Patterned	Short	Tiny	Tiger	Wild
46	Low	High	Patterned	Medium	Elongated	Tiger	Wild
47	Average	High	Patterned	Short	Tiny	Cheetah	Wild
48	High	Low	Patterned	Medium	Elongated	Cheetah	Wild
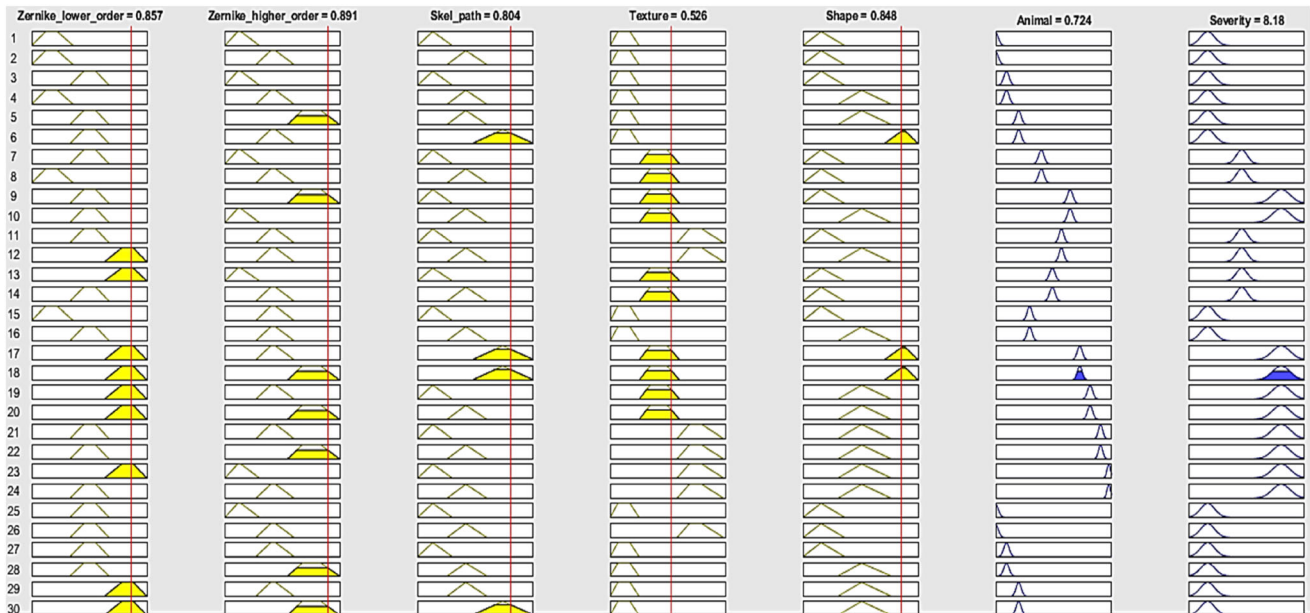
Fig. 8 Defuzzification of elephant with the category as wild animal

Fig. 9 Sample images from thermal dataset

Fig. 10 Challenging image conditions

whereas Zernike and shape features have achieved considerable performance and proved to be the foremost features necessary for accurate recognition of animals. Nevertheless, the fusion of all four feature descriptors has the highest accuracy of 97% implying the necessity of all four feature sets. The best results are highlighted in bold-face from Table 3 to Table 5.

One of the common issues in camera trap images is animals being captured too far away or too close to the camera. To study this issue, we deliberately captured the animals from varying distance in the distance range of 5 to 50 m. The effect of distance from thermal camera is illustrated in Fig. 12.

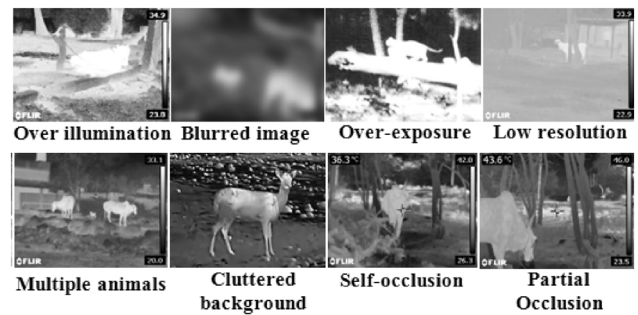The number of incorrect results was quite high for monkeys, as they were mostly captured within dense vegetation. Hence, distance solely does not affect the recognition rate, but distance coupled with vegetation affects the recognition. In a yet another experiment, the animal detection from challenging images was studied with both volunteers and IF-FUZ (see Fig. 13).

IF-FUZ detected animals accurately from almost all the challenging images, except in self-occluded and partially captured images, whereas volunteers found it difficult to detect animals from images having illumination and resolution problem.

3.3 Comparative Study

For the baseline comparison, we compared IF-FUZ with other fuzzy models on our thermal dataset (see Table 4). Among all, the Neuro-fuzzy models [15, 16] achieved substantial performance when compared to other systems. However, with carefully chosen invariant features, IF-FUZ achieved the best results.
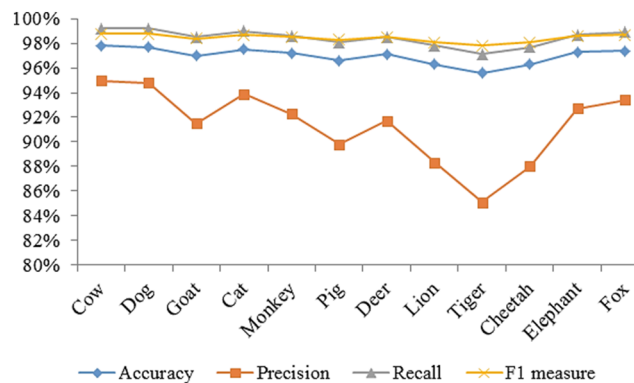
Fig. 11 Performance analysis of IF-FUZ on thermal dataset

Table 3 Influence of discrete features and their fusion

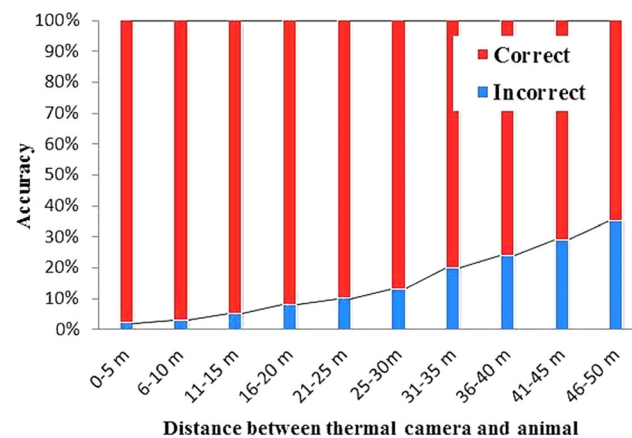| Discrete features	Accuracy (%)
LBP	65.39
Zernike	84.91
Skeleton	72.96
SMM-MAM	81.48
Feature fusion	
Zernike + LBP	81.45
Zernike + skeleton	77.12
Zernike + SMM-MAM	88.93
LBP + skeleton	71.69
LBP + SMM-MAM	84.92
Skeleton + SMM-MAM	73.41
Zernike + LBP + skeleton + SMM-MAM (IF-FUZ)	**97.0**

Fig. 12 Effects of distance in recognizing animals

Camera traps images have several challenging image conditions that make the animal detection a difficult task. Here, we have represented 10 most common challenges in camera traps and identified the best feature for each of the
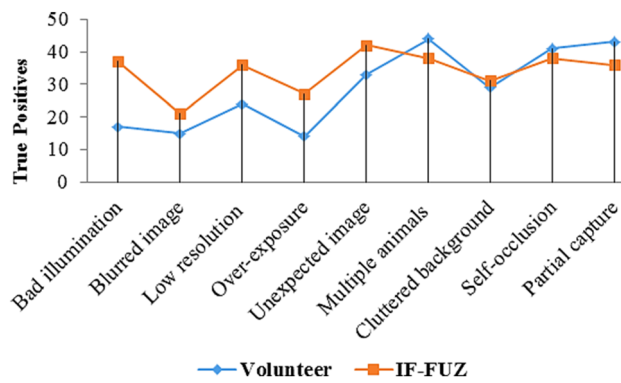

Fig. 13 Performance comparison in detecting animals from challenging images

Table 4 Comparison with existing fuzzy models

Refs.	Method	Accuracy (%)	PPV (%)	TPR (%)
[12]	FL-TM	79.3	83	78
[13]	F-FNN	89.4	87	77
[15]	FCM-CNN	85.4	84	85
[16]	FIS-CNN	91.1	80	87
[17]	FL-IP	86.6	86	71
Ours	IF-FUZ	**97.0**	**91**	**98**

challenges. We also compared the performance of animal–vehicle detection systems [6, 7] with IF-FUZ on the thermal dataset and the result is presented in Table 5.

As inferred from the table, for challenges that arises due to lighting (illumination problem, over-exposure, or low resolution images), the illumination invariant LBP feature was most appropriate in detecting the animal. Nevertheless, the fusion of features has higher confidence in the detection. We also compared the performance of other related systems [6, 7] in detecting animals from challenging images. A sample image for each challenging condition in Table 5 is represented in Fig. 10. We also studied how far animal detection is efficient with thermal and RGB images. As part of the study, we compared IF-FUZ with RGB-based animal detection systems [18, 22] and the result is depicted in Fig. 14.

As inferred from the figure, the animal detection was lot easier and accurate with thermal images than RGB, although the contrast between animal and the background in the thermal images is poor. This signifies that thermal images are most appropriate for capturing animals than RGB. Among the compared methods, IF-FUZ had accurate results, whereas the compared systems produced FPs and FNs, due to the choice of image modal.

Table 5 Accuracy of discrete features in detecting animals from challenging images

| Challenging characteristics	Accuracy of IF-FUZ (%)					Existing systems (%)	
	Zernike	Skeleton	LBP	Shape	Fusion	[6]	[7]
Poor illumination	77.65	71.77	**89.35**	66.75	90.87	85.87	74.65
Over illumination	70.14	84.25	**88.70**	79.51	91.56	83.56	61.14
Blurred image	64.66	65.46	**77.27**	72.59	86.43	77.43	59.66
Low resolution	58.79	69.59	**74.55**	74.37	84.96	77.96	51.79
Over-exposure	86.33	72.86	**97.59**	73.93	90.58	79.58	78.33
Unexpected image	**98.12**	87.32	62.24	93.87	88.32	79.32	89.12
Multiple animals	**95.86**	92.35	64.26	90.33	97.35	88.35	83.86
Cluttered background	56.32	62.43	54.47	**79.50**	90.83	87.83	46.32
Self-occlusion	70.27	88.67	65.48	**88.73**	96.67	86.67	59.27
Partial occlusion	64.75	**80.68**	59.86	75.27	91.96	88.96	55.75
Partial capture	73.86	**94.80**	93.27	74.70	96.67	84.67	68.86
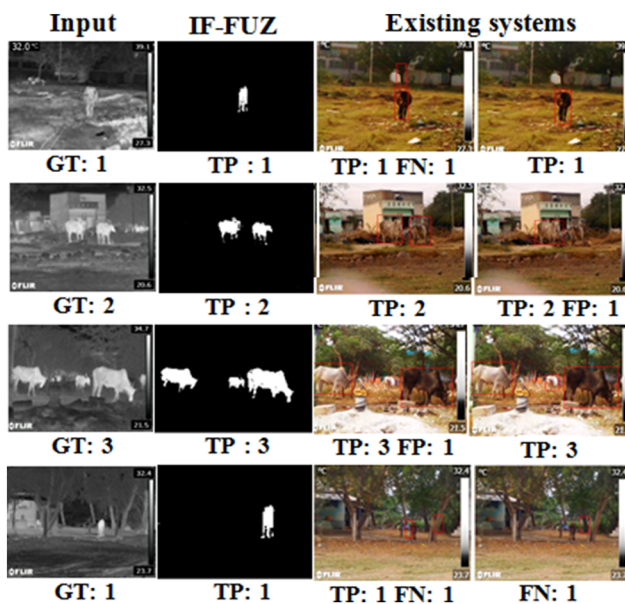
Fig. 14 Animal detection in RGB vs. Thermal images

3.4 Discussion

The results from the experiments validate the suitability of thermal images for animal-based applications. Animals are different from humans and the model should be designed catering to the characteristics of animals. Evidently, invariant features effectively detect and recognize the animals even from challenging image conditions. Hence, the features should be chosen to adapt these characteristics of animals. With thermal images, we can ignore the camouflage due to dense vegetation. Although the model had good precision, the detection was quite difficult during day time as the surroundings get heated up due to the soaring temperature and it becomes difficult in differentiating the foreground and background in certain images. For such images, the pre-processing had little effect and so we

conclude that thermal images are most useful during pre-dawn, dusk and night time. It is also worth noting that resolution decreases with broader FOV. The proposed model is complementary to other animal warning systems and IF-FUZ can be used as a non-audible alert to forest officials to take necessary actions. Fuzzy constraints [23] and Neuro-fuzzy techniques [24] can be combined to make further intelligent models. Similarly, thermal and visible images [25] can be fused to obtain complementary information from both the image modalities and can overcome the drawbacks from individual images. Besides, thermal images can be employed in aerial animal detection applications [26] for monitoring animals over a vast area.

4 Conclusion and Future Scope

In this work, we proposed a robust animal detection and recognition system using fuzzy logic and thermal images. The proposed model entitled IF-FUZ encompasses the most appropriate invariant features, without overloading the fuzzy inference system. Through experimental results it is inferred that the invariant features used in the fuzzy system are efficient in detecting animals even from the most challenging image conditions. Under most of the circumstances, the detection rate was close to 100% and the recognition rate was around 98%. Besides, the proposed model holds good for any animal-oriented applications. As part of the future work, the model has to be fine-tuned to detect animals from RGB images as well, as they can be helpful during day time. Alternatively, fusion of visible and thermal images can be studied for detecting animals. In addition, the model has to be tested under various altitude and vegetation, to better understand the effects of thermal imaging.

References

1. Ganow, K.B., Caire, W., Matlack, R.S.: Use of thermal imaging to estimate the population sizes of Brazilian free-tailed bat, Tadarida Brasiliensis, maternity roosts in Oklahoma. Southwestern Nat. **60**(1), 90–96 (2015). https://doi.org/10.1894/SWNAT-D-14-00010R1.1

2. Hristov, N.I., Betke, M., Theriault, D.E., Bagchi, A., Kunz, T.H.: Seasonal variation in colony size of Brazilian free-tailed bats at Carlsbad Cavern based on thermal imaging. J. Mammal. **91**(1), 183–192 (2010). https://doi.org/10.1644/08-MAMM-A-391R.1

3. Welbourne, D.U.S.T.E.N.: A method for surveying diurnal terrestrial reptiles with passive infrared automatically triggered cameras. PLoS ONE **6**, e18965 (2013)

4. Goodenough, A.E., Carpenter, W.S., MacTavish, L., MacTavish, D., Theron, C., Hart, A.G.: Empirically testing the effectiveness of thermal imaging as a tool for identification of large mammals in the African bushveldt. Afr. J. Ecol. **56**(1), 51–62 (2018). https://doi.org/10.1111/aje.12416

5. Barbosa Pereira, C., Kunczik, J., Zieglowski, L., Tolba, R., Abdelrahman, A., Zechner, D., Czaplik, M.: Remote welfare monitoring of rodents using thermal imaging. Sensors **18**(11), 3653 (2018). https://doi.org/10.3390/s18113653

6. Zhou, D., Dillon, M., Kwon, E.: Tracking-based deer vehicle collision detection using thermal imaging. In: 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO) (pp. 688–693). IEEE. (2009) https://doi.org/10.1109/robio.2009.5420589

7. Zhou, D., Wang, J., Wang, S.: Contour based HOG deer detection in thermal images for traffic safety. In: Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV) (p. 1). The steering committee of the world congress in computer science, computer engineering and applied computing (WorldComp) (2012)

8. Cilulko, J., Janiszewski, P., Bogdaszewski, M., Szczygielska, E.: Infrared thermal imaging in studies of wild animals. Eur. J. Wildl. Res. **59**(1), 17–23 (2013). https://doi.org/10.1007/s10344-012-0688-1

9. Forslund, D., Bjärkefur, J.: Night vision animal detection. In: 2014 IEEE intelligent vehicles symposium proceedings (pp. 737–742). IEEE. (2014) https://doi.org/10.1109/ivs.2014.6856446

10. Roy, S., Shivakumara, P., Jain, N., Khare, V., Dutta, A., Pal, U., Lu, T.: Rough-fuzzy based scene categorization for text detection and recognition in video. Pattern Recogn. **80**, 64–82 (2018). https://doi.org/10.1016/j.patcog.2018.02.014

11. Darwich, A., Hébert, P.A., Bigand, A., Mohanna, Y.: Background subtraction based on a new fuzzy mixture of Gaussians for moving object detection. J. Imaging **4**(7), 92 (2018). https://doi.org/10.3390/jimaging4070092

12. Mahapatra, A., Mishra, T. K., Sa, P. K., Majhi, B.: Background subtraction and human detection in outdoor videos using fuzzy logic. In: 2013 IEEE international conference on fuzzy systems (FUZZIEEE) (pp. 1–7). IEEE. (2013) https://doi.org/10.1109/fuzz-ieee.2013.6622397

13. Toran, V., Sipi, D.: Fuzzy-filtered neural network for rice disease diagnosis using image analysis. Int. J. Innov. Technol. Explor. Eng. **8**, 437–446 (2019)

14. Brunassi, L.D.A., Moura, D.J.D., Nääs, I.D.A., Vale, M.M.D., Souza, S.R.L.D., Lima, K.A.O.D., et al.: Improving detection of dairy cow estrus using fuzzy logic. Scientia Agricola **67**(5), 503–509 (2010). https://doi.org/10.1590/S0103-90162010000500002

15. John, V., Mita, S., Liu, Z., Qi, B.: . Pedestrian detection in thermal images using adaptive fuzzy C-means clustering and convolutional neural networks. In: 2015 14th IAPR international conference on machine vision applications (MVA) (pp. 246–249). IEEE. (2015) https://doi.org/10.1109/mva.2015.7153177

16. Kang, J.K., Hong, H.G., Park, K.R.: Pedestrian detection based on adaptive selection of visible light or far-infrared light camera image by fuzzy inference system and convolutional neural network-based verification. Sensors **17**(7), 1598 (2017). https://doi.org/10.3390/s17071598

17. Jain, I., Rani, B.: Vehicle detection using image processing and fuzzy logic. Int. J. Comput. Sci. Commun. **1**(2), 255–257 (2010)

18. Meena, S.D., Agilandeeswari, L.: Stacked convolutional autoencoder for detecting animal images in cluttered scenes with a novel feature extraction framework. Soft computing for problem solving, pp. 513–522. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-0184-5_44

19. Stubendek, A., Karacs, K.: Shape recognition based on projected edges and global statistical features. Math Prob Eng (2018). https://doi.org/10.1155/2018/4763050

20. F. systems. Thermal imaging, night vision and infrared camera system. https://www.flir.com. Accessed 3 Jan 2019

21. Meena, S.D., Agilandeeswari, L.: An Efficient framework for animal breeds classification using semi-supervised learning and Multi-Part Convolutional Neural Network (MP-CNN). IEEE Access **7**, 151783–151802 (2019). https://doi.org/10.1109/ACCESS.2019.2947717

22. Meena, S.D., Agilandeeswari, L.: Adaboost cascade classifier for classification and identification of wild animals using movidius neural compute stick. Int. J. Eng. Adv. Technol. **9**(13), 495–499 (2019). https://doi.org/10.35940/ijeat.a1089.1291s319

23. Sun, K., Mou, S., Qiu, J., Wang, T., Gao, H.: Adaptive fuzzy control for non-triangular structural stochastic switched nonlinear systems with full state constraints. IEEE T Fuzzy Syst **27**, 1587–1601 (2018). https://doi.org/10.1109/TFUZZ.2018.2883374

24. Sun, K., Qiu, J., Karimi, H.R., Gao, H.: A novel FiniteTime control for nonstrict feedback saturated nonlinear systems with tracking error constraint. IEEE Trans. Syst. Man Cybern. Syst. (2019). https://doi.org/10.1109/tsmc.2019.2958072

25. Divya, M.S., Agilandeeswari, L.: A new supervised clustering framework using multi discriminative parts and expectation–maximization approach for a fine-grained animal breed classification (SC-MPEM). Neural Process. Lett. (2020). https://doi.org/10.1007/s11063-020-10246-3

26. Divya, M.S., Agilandeeswari, L.: FSSCaps-DetCountNet: fuzzy soft sets and CapsNet-based detection and counting network for monitoring animals from aerial images. J. Appl. Remote Sens. **14**(2), 026521 (2020). https://doi.org/10.1117/1.JRS.14.026521

S. Divya Meena received the B.Tech degree in Information technology from the Vellore Institute of Technology (VIT), Vellore, India, in 2014, and the M.E. degree in Computer science and engineering from Anna University, Chennai, in 2016. She is currently pursuing the Ph.D. degree with the School of Information Technology and Engineering, VIT. From 2016 to 2017, she was an Assistant Professor with the Jansons Institute of Technology, Coimbatore, India. Her research interests include deep learning, thermal imaging, fuzzy logic and remote sensing.

L. Agilandeeswari received the bachelor's degree in Information Technology and the master's degree in Computer Science and Engineering from Anna University, in 2005 and 2009, respectively, and also completed the Ph.D. degree. She is currently the HOD and an Associate Professor with the Department of Digital Communications, School of Information Technology and Engineering (SITE), Vellore Institute of Technology (VIT), Vellore. She is having around 13+ years of teaching experience. She has published 50+ articles in peer-reviewed reputed journals and 13 engineering books. Her reputed publications include research articles in peer-reviewed journals namely Expert Systems with Applications, IEEE Access, Journal of Ambient Intelligence and Humanized Computing, Multimedia Tools and Applications, Journal of Applied Remote Sensing indexing at Thomson Reuters with average impact factor of 5. She is the peer reviewer in journals such as IEEE Access, International Journal of Remote Sensing, Array, Artificial Intelligence Review, Informatics in Medicine Unlocked, Neuro computing, Computers and Electrical Engineering, Journal of King Saud University—Computer and Information Sciences, IET ReView, Journal of Engineering Science and Technology (JESTEC), etc. She is a Life Time Member of the Computer Society of India. Her areas of interests include Image and Video Watermarking, Image and Video Processing, Machine Learning, Fuzzy logic, Cryptography, IoT, Information Centric Networks, and remote sensing. She has also got the Best Researcher Award for 5 years at SITE, VIT Vellore, from 2015 to 2019.