Journal of Theoretical and Applied Information Technology
15th May 2017. Vol.95. No 9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMPUTER VISION BASED DANCE POSTURE EXTRACTION USING SLIC

K.V.V. KUMAR[1], P.V.V. KISHORE[2], A.S.C.S. SASTRY[3], D. ANIL KUMAR[4], E.KIRAN KUMAR[5]

[1, 2,3,4,5] Department of Electronics and Communications Engineering, K L University, Green Fields,

Vaddeswaram, Guntur DT, Andhra Pradesh, India.

E-mail: kumarece405@gmail.com, pvvkishore@kluniversity.in, ascssastry@kluniversity.in, danilmurali@kluniversity.in, kiraneepuri@kluniversity.in

ABSTRACT

Indian Cultural dance poses are segmented using two algorithms and their performance is estimated quantitatively. A data set of various dance poses are captured under controlled environments. The image dataset consists of 100 dance mudras of kuchipudi dance poses. Marker controlled Watershed and super pixel based simple linear Iterative clustering (SLIC) algorithms are proposed to accomplish the segmentation task. The results of both the algorithms show that SLIC based super pixel segmentation outperforms the watershed algorithm. Further, the results of the segmentation can be used for classification of various Indian dance forms.

Keywords: *Dance Image Segmentation, Watershed Algorithm, Super Pixels, Simple Linear Iterative Clustering.*

1. INTRODUCTION

Indian dance forms are a mirror to rich cultural heritage that existed from the past 5000 years. The name for these classical dance forms is called 'Natya Rasa' as portrayed in the bible of Indian dance 'Natya Shastra' According to Natya Shastra there are 108 karanas [1] meaning action of hands, feet and body. These poses symbolize various physical meaning related to nature, god and actions. A few hasta mudras are shown in figure 1 for reference [2].

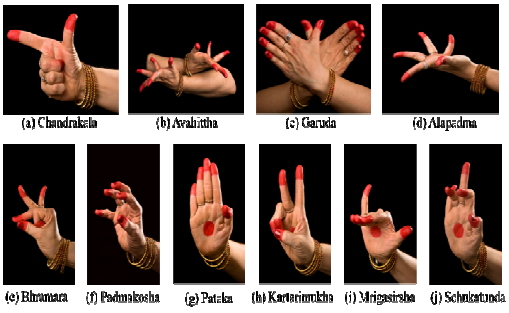

Figure.1. Hasta Mudras of Indian Classical Dances

Gesture based communication is a Computer vision based dialect for hard of hearing and listening to individuals, which includes utilizing hands, face and body. Present work complex in constant communication through sign video postures issue in Computer vision. Numerous works center in the cluster gives the gesture based communication acknowledgment.

The accompanying framework includes pre-processing, division, highlight extraction and acknowledgment frameworks. The main parameters are hand movements and head area in video outlines. This work looks at pictures and constant Indian gesture based communications recordings under ordinary examples with simple foundations by using SLIC (Simple Linear Iterative Clustering).

Super pixel calculations social occasions of pixels into perceptually basic nuclear regions which can be utilized to supplant the dissipating structure of the pixel grid. They distinguish picture wealth, give a productive primitive from which to process picture highlights, and through and through diminish the disperse way of resulting picture get prepared assignments. They have possessed the capacity to be key building squares of different PC vision figuring, for occurrence, top scoring multiclass object division portions. Challenge, centrality estimation, division, body model estimation, and thing restraint. There are different ways to deal with oversee conveying super pixels, each with its own reasons for side interest and weights that might be more qualified to a specific application. For instance, if adherence quite far is of key monstrosity, the framework based procedure of might be a perfect decision. In any case, if super pixels are to be used to produce a diagram, a framework that makes a more standard cross range,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

for occasion, is likely an overwhelming decision. While it is troublesome to portray what constitutes a flawless system for all applications, we are using the going with properties which are alluring:

1.Super pixels should stick well to limits of picture.

2.Now that used to diminish computational disperse quality as a pre-processing step, super pixels should rush to enroll, memory successful, and simple to use.

3.Now that used for division purposes, super pixels should both developments the speed and upgrade the way of the picture.

The idea is to represent Indian classical dance on a digital platform. Indian cultural dance forms are most complex human gestures to be represented in digital format. Feature extraction is most complicated task as the images are full of color, occlusions and finger closeness. This can be observed in images in figure 1. The regular image processing segmentation models such as thresholding and edge detection fail to represent the correct shapes as found in the original images.

2. LITERATURE SURVEY

The idea is to represent Indian classical dance on a digital platform. Indian cultural dance forms are most complex human gestures to be represented in digital format. Feature extraction is most complicated task as the images are full of color, occlusions and finger closeness. The regular image processing segmentation models such as thresholding and edge detection fail to represent the correct shapes as found in the original images.

The most vibrant and highly used models of segmentation in recent times are active contours. Many models have been proposed in literature [3] [4]. But the basic model sufferers from many drawbacks such as illumination, position of the mask and number of iterations. We believe focused active contour models with more spatial information using color, texture and shape have profound effect on extracting the correct segments [5].

The procedure of isolating a picture into more number of small parts in which the cluster or gathering of pixels with a specific end goal to identify the picture is called as image segmentation. It also changes the representation of a photo into something that is more analyzed and less requesting to separate is the essential point or target of division. discovering objects and boundaries (lines, ridges, edges etc. Here each pixel is compared with the new pixel so that the total distance is calculated. The distance between the pixel which are having same characteristics we are having color special and having color position.

The image is divided within the sequence of combinations of the portions within the entire picture and during the specification of pixels is done within the shapes. A part of the values or figured property (such as shading, power, or surface) are spoken to in each of the pixel in a location, neighboring locales or pixels are distinctive regarding the same characteristics. The shapes after picture division is not therapeutic imaging can be utilized to make three-dimensional recreations with the assistance of insertion calculations like coordinating blocks.

The division is an imperative stage in pictures and translation preparing. There are two fundamental ways to deal with division: the boondocks approach and the district approach. The division by watershed join the two methodologies. This is an intense system for fast location of both edges and locales. The real issue of the watershed change is over-division. Undoubtedly, this calculation is delicate to any neighborhood least in the picture, and tends to characterize the lines of the watershed change where every nearby least offers ascend to an area. To stay away from this issue, effective instruments adjusted to various issues have been proposed in the writing. Either decrease the quantity of minima and poster estimation of too numerous districts. Proceed by either separating strategies by combining the districts as indicated by comparability criteria after phantom and spatial use of the watershed.

To address this, we propose another super pixel calculation: basic Simple Linear Iterative Clustering (SLIC), which adjusts k-implies grouping to produce super pixels in a way like. While strikingly basic, SLIC is appeared to yield best in class adherence to picture limits on the Berkeley benchmark, and having existing techniques when utilized for division on the PASCAL what's more, MSRC information sets. Moreover, it is quicker and more memory productive than existing strategies. Here not withstanding these quantifiable advantages, SLIC is anything but difficult to utilize, offers adaptability in the smallness what's more, number of the super pixels it creates, is clear to reach out to higher measurements, and is openly accessible.

3. METHODOLOGY

3.1 Dataset for Kuchipudi Dance Form

The dataset is made from a combination of lab dance mudras and dance mudras images on the websites of Indian art and culture. For each mudra

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

we have made a set of 5 images from 5 different artists. Figure 2 shows the set used for capturing the dataset at university multimedia centr.

A mixture of 5×24 images is used for training and testing. We have 2 sets of images from our dancers, 2 sets from dance websites and 1 set from YouTube video frames.

Figure 2. Setup Used For Capturing Dance Images Of Kuchipudi Mudras

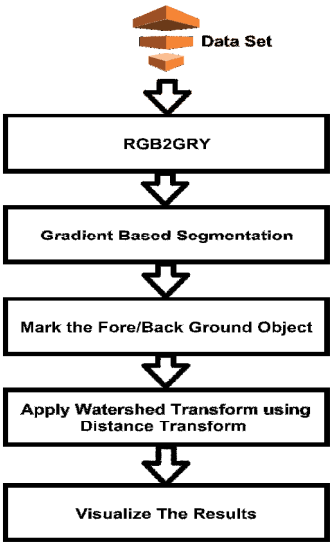

Figure3. Flow Chart for Watershed Segmentation

3.2 Watershed Algorithm

Watershed change is an approach to confining covering objects. A watershed [6] is molded by "flooding" a photo from its neighborhood minima and encircling "dams" where waterfronts meet. Right when the photo is totally flooded, all dams together shape the watershed of a photo. For division of a photo ought to be conceivable with watershed (of the primary picture). The catchment dishes of the watershed contrast with articles, while envisioning the edginess picture in three-dimensional scene i.e., the watershed of the edgeless picture will exhibit as far as possible.

Consistently, some pre-and post-get ready of the watershed picture is relied upon to procure a not too bad division. Filling in of the points of confinement to obtain solid parts is a post taking care of. By watershed division, little breaks in a point of confinement may not be concerned. Because of the way the watershed is fabricated, spills can't happen. Spills can't happen because of the way the watershed is manufactured.

To avoid an over division it is vital to realize some pre and post taking care of. In our example, each dull quality is set underneath a breaking point to zero. This promises no dams are formed in the early flooding process. Right when such dams are brought on by fragile edges, they don't identify with inquiry constrains and are not a bit of watershed division. Some over division is showed up nearby the right question limits in the example. Another method that is intense to the delineation is to learn the typical edginess estimation of each branch and remove those branches if ordinary falls underneath utmost**.**

The watershed change is a device morphological based for picture division. It is the method which consider a dark level picture as a topographic alleviation. If one consolidates the dim level of every point at a height. It is then conceivable to characterize the watershed change as the edge framing the limit between two watersheds. This is to register the watershed of the said alleviation. Watersheds in this way got relate to areas of the picture. Watershed speaks to the limits between neighboring catchments. The base can be deciphered as markers of watershed districts and the watershed can be translated as forms.

3.2.1. Thresholding

Thresholding is the most straightforward technique for picture division. This thresholding image is used to change the gray scale image to the matched image so that matched pair of the image is given. It is used to separate the image into the RGB

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

scale in order to good image with better thresholding value.

The leading role to choose the correct pixel value with the estimated value which gives the better thresholding value so that better decision is taken and which is compared with other images (or values when numerous levels are chosen). A few techniques are utilized for industry including the greatest entropy strategy, yield strategy (most extreme change), and k-implies clustering.

As of late, a few techniques for thresholding Computed Tomography (CT) pictures have been produced. This threshold images which are similar to the normal images so that the better images are given. It is compared in order to the noisy between the pixels.

There are many techniques so that the single dimensional images are taken from the other source so that the better results are produced. Some thresholding is set up in order to convert the color image into the gray scale image. These method for the images which have noisy between the pixels so that the better image is present and to good pixel values overall every fragment of total performance.

3.2.2. Clustering

It is a repetitive procedure to perform the number of iterations in which the image is converted to sub-images in which the total pixels are taken.

1. Initially number of cluster S are taken so that the number of cluster centers are found.
2. Taking that each every is free from noisy pixel adjoining with other pixel with surrounding pixels.
3. The following process is repeated to produce the results with positive gradient position.
4. The step2 and step 3 are done iteratively to produce the least distance.

For this circumstance, the partition between a pixel and a group center is squared or add up to balance is figured with relies on upon pixel shading, power, creation and zone or a weighted mix of these segments. S can be picked physically, subjectively, or by a heuristic. Serious the algorithm is guaranteed to join, it may not give back the perfect course of action. Starting course of action of clusters and the estimation of K describes the way of game plan.

Clustering and arrangement are both essential errands in signal processing which plays a key role and gives idea of centering of a pixel. It is used generally as an oversaw learning technique, gathering for unsupervised adjusting some bunching models are for both. The target of bunching is unmistakable, that of request is farsighted. Since the target of grouping is to discover another plan of arrangements, the new social occasions are of eagerness for themselves, and their assessment is characteristic. All together assignments, nevertheless, an essential part of the examination is unessential. Since the social events must reflect some reference set of classes.

Clustering groups information cases into subsets in such a way, to the point that comparable occurrences are gathered together, while diverse cases have a place with various groups the events are along these lines sorted out into a profitable representation that depicts the masses being tried. Formally, the bunching structure is addressed as a plan of subsets K = K1; K2; … ...Kk of 'L', such that: L =Lk, z=C_z and z for r=t Here, any event in S has a spot with definitely one and one and just subset. Grouping of articles is as old as the human necessity for depicting the striking characteristics of men and addresses and remembering them with a sort. In this manner, it handles distinctive legitimate controls: from number-crunching and estimations to science and genetic qualities, each of which uses different terms to delineate the topologies surrounded using this examination. From normal "logical arrangements", to therapeutic "issue" and innate "genotypes" to gathering "bunch advancement" the issue is indistinguishable: shaping classifications of elements and doling out people to the best possible gatherings inside of it.

3.2.2.1 Types of clustering

Around there we portray the most without a doubt comprehended gathering estimations. The maximum difference in which the clustering systems and various clusters which is not equivocally given in the system. Consequently, various clustering techniques have been made, each of which uses another insincerity standard Bundling propose isolating the gathering methods into two guideline groups: different levelled and distributing. It prescribes ordering the procedures into additional three crucial orders: thickness based methodologies, model-based grouping and matrix based strategies. An alternative game plan considering the activation principle of the diverse packing procedures is presented.

1.Graph-Theoretic Clustering

The occasions spoke to as hubs. A surely understood diagram theoretic calculation is in light of the Minimal Spanning Trees. It is used to find the edges which are normal to the edges within the charts and which are adjacent to the neighborhood

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

sets. There we are having the connections between the levelled techniques and different techniques related to the theoretic clustering between the clusters.

2. Single-connection clusters

There are the graphs used in this network so that the subgraph is checked with the large graph. The graph is defined as the specific work within less individual work from the set so that the total set is taken from the big positon and minimum distance is taken within the least distance.

3. Complete-join bunches

Here we are having the clusters and we are having the limit values so that the total image is taken. Every cluster with the less samples and more joints the cluster is taken and so that the maximum value is taken with concerning following property.

3.3 SLIC Super-Pixel Segmentation

The super pixels [7,8,9] contrast with gatherings in the rvwxy shading picture plane space. This shows an issue in portraying the division measure D, which may not be immediately undeniable. D enrols the partition between a pixel 'i' and gathering center Lk in Algorithm.

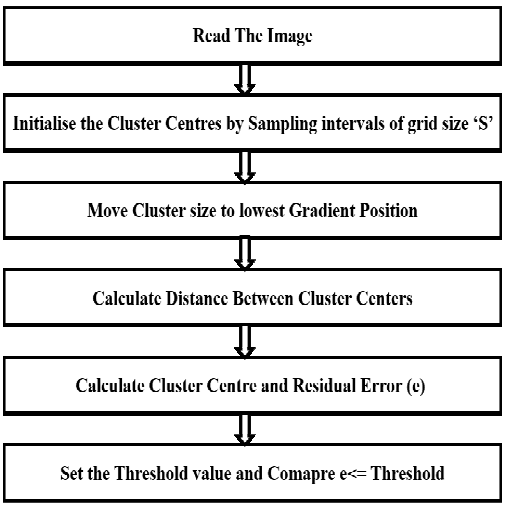

Figure4. Flow Chart for SLIC Segmentation

A pixel's shading is addressed in the shading space whose extent of possible qualities is known. The pixel's position, of course, may take an extent of characteristics that contrasts as demonstrated by the range of the photo not hold quick well as far as possible. For humbler super pixels, the visit is legitimate. To solidify the two partitions into a single measure, it is imperative to institutionalize shading region and spatial proximity by their specific most compelling detachments within a gathering. Doing accordingly, D0 is formed.

$$l_m = \sqrt{((v_j - v_i)^2 + (r_j - r_i)^2 + (w_j - w_i)^2)} \quad (1)$$

$$l_j = \sqrt{((y_j - y_i)^2 + (z_j - z_i)^2)} \quad (2)$$

$$L' = \sqrt{\left(\left(\frac{l_m}{h_c}\right)^2 + \left(\frac{l_j}{h_s}\right)^2\right)} \quad (3)$$

The most extreme spatial separation expected inside of a given group should compare to the testing interim, $K_l = K = \sqrt{\left(\frac{L}{M}\right)}$ and the N_c which the cluster center is taken. The distance is taken and in which the cluster centers are taken and square root of the gradient position. It gives the positive magnitude in which the square of the field. The distance is used to calculated by using equation 3.

Here we are going to characterized some of the D values in which the significance between the gray scale image is done. So, that the total image redundancy is reduced by the normal distances. At particular point in which the big distances are calculated between the matched point and fluffy image so that the total distance is found by using the same characteristics of the pixels so that same types of images are taken. The grayscale and color spatial images are taken to detect the images using some of the methods.

$$k_i = \sqrt{((f_m - f_n)^2)} \quad (4)$$

It is likely same type of distance so that Hough man transform images is done so that the total distance is formed.

$$H_k = \sqrt{((v_m - v_n)^2 + (r_m - r_n)^2 + (w_m - w_n)^2)} \quad (5)$$

3. RESULTS AND DISCUSSION

The dataset consists of 100×5 images of Indian dance form kuchipudi are collected from various sources. These 500 images are contrast enhanced by 20% to smooth pixel values. We came to understand that these basic mudras are common to all Indian dance forms listed in [1]. Watershed and Super pixel based segmentation algorithms are initiated on the data set.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

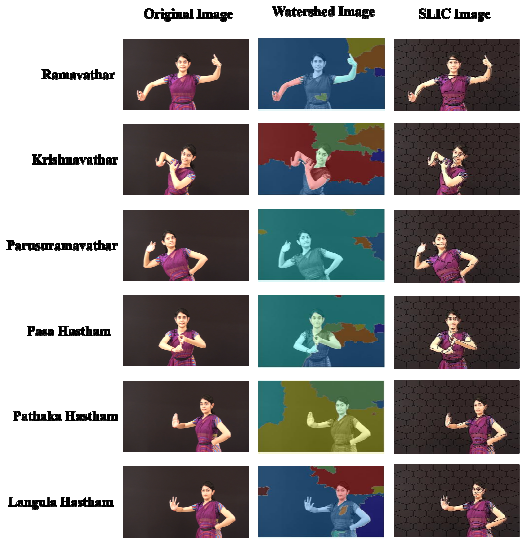

Figure 5: Results of the proposed algorithms

Figure 5 shows the results of segmentation of both algorithms. Visual quantification of the results from the figure 5 shows that the SLIC based super pixel can segment each portion of the dancer to near perfection compared to marker controlled watershed algorithm.

Watershed algorithm gives various colors to segmented parts. These parts should collectively represent a dance pose. The parts expected from watershed algorithm represent an average of 52.55% of the total images. In case of Super pixel based SLIC the percentage is around 74.25%. These percentages are calculated based on the ground truth models generated manually.

The other problem encountered is in the number of segments produced per image mudra. Depending on the hand density in the image frame, the number of segments in each feature vector are of different size. To tackle this problem, we modified the segments of all images to a normalized size. The normalization process involved selection of important segments based on magnitude of feature vectors in the particular segments.

Performance evaluation between the proposed methods for dance pose segmentation is done using Structural Similarity Index measure (SSIM) [10], peak signal to noise ratio(PSNR) [10] and image quality index(IQI) [10].

Table 1 presents values of SSIM [11] for different mudras for the watershed and SLIC base pose segmentation. The SLIC algorithm shows a high value of SSIM compared to watershed.

Table 2 presents values of PSNR for different mudras for the watershed and SLIC base pose segmentation. The SLIC algorithm shows a high value of PSNR compared to watershed.

Table 3 presents values of IQI for different mudras for the watershed and SLIC base pose segmentation. The SLIC algorithm shows a high value of IQI compared to watershed.

Table I. SSIM Based Comparison

FRAME NO	WATERSHED BASED POSE ESTIMATION	SLIC BASESD DANCE POSE ESTIMATION
1	**0.768**	**0.871**
2	**0.781**	**0.882**
3	**0.756**	**0.848**
4	**0.749**	**0.857**
5	**0.752**	**0.887**
6	**0.759**	**0.862**
7	**0.761**	**0.851**
15	**0.774**	**0.879**
16	**0.778**	**0.873**
17	**0.761**	**0.871**
18	**0.401**	**0.581**
19	**0.424**	**0.596**
35	**0.398**	**0.571**
36	**0.381**	**0.554**
37	**0.427**	**0.607**
38	**0.411**	**0.594**
39	**0.394**	**0.569**
40	**0.417**	**0.599**

Table II. PSNR Based Comparison

FRAME NO	WATERSHED BASED POSE ESTIMATION	SLIC BASESD DANCE POSE ESTIMATION
1	**9.15**	**12.74**
2	**9.27**	**12.59**
3	**8.81**	**12.46**
4	**9.26**	**12.57**
5	**9.38**	**12.49**
6	**9.29**	**12.65**
7	**8.92**	**12.38**
15	**9.09**	**12.29**
16	**9.21**	**12.56**
17	**9.51**	**12.93**
18	**6.91**	**10.12**
19	**7.01**	**10.26**
35	**6.96**	**10.04**
36	**6.89**	**9.98**
37	**7.12**	**10.46**
38	**6.99**	**10.01**

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

39	**7.14**	**10.17**
40	**7.07**	**10.31**

Table III. IQI Based Comparison

FRAME NO	WATERSHED BASED POSE ESTIMATION	SLIC BASESD DANCE POSE ESTIMATION
1	**0.771**	**0.876**
2	**0.774**	**0.872**
3	**0.761**	**0.851**
4	**0.746**	**0.864**
5	**0.759**	**0.891**
6	**0.763**	**0.861**
7	**0.752**	**0.858**
15	**0.781**	**0.877**
16	**0.773**	**0.876**
17	**0.778**	**0.881**
18	**0.405**	**0.584**
19	**0.431**	**0.605**
35	**0.401**	**0.569**
36	**0.385**	**0.563**
37	**0.431**	**0.614**
38	**0.403**	**0.602**
39	**0.405**	**0.572**
40	**0.424**	**0.611**

4. CONCLUSION

An attempt is made to segment dance mudras of Indian classical dance from kuchipudi based on image processing models. Two segmentation extraction techniques are compared for this work. Watershed and SLIC super pixel algorithms are tested on a dataset of 500 dance images with 100 mudras. Visual verification and structural verification using SSIM are preformed to check the classifiers performance. The Super pixel based SLIC segregation registered an average segmentation score of 74.25%. Most of the mudras with two hands and head produced occlusions that induced bottleneck's during feature extraction stage. This model of mudra classification will help enhance the learning capacity of a first-time learner.

REFRENCES:

[1] Naidu, N. Naidu, B.V. Pantulu, P.R. Tandavan Lakshanam, A Karana in dance is defined as "the coordination of the movements of the hands and feet", *The Fundamentals of Ancient Hindu Dancing*. New Delhi, Munshiram Manoharlal, 1971, 1980, pp.19.

[2] http://natyanjali.blogspot.in/

[3] Kishore, P. V. V., S. R. C. Kishore, and M. V. D. Prasad. "Conglomeration of hand shapes and texture information for recognizing gestures of Indian sign language using feed forward neural networks." *International Journal of engineering and Technology (IJET), ISSN* (2013): pp.0975-4024.

[4] Anandh, A., K. Mala, and S. Suganya. "Content based image retrieval system based on semantic information using color, texture and shape features." In *Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on*, pp. 1-8. IEEE, 2016.

[5] Chen, Shiyu, et al. "poor textural image matching based on graph theory." International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences 41 (2016).

[6] Kishore, p.v.v., and Madhav, B.T.P., "flower image segmentation: a comparison between watershed, marker controlled watershed, and watershed edge wavelet fusion." (2016).

[7] A. Lucchi, K. Smith, R. Achanta, G. Knott, and P. Fua, "Supervoxel-Based Segmentation of Mitochondria in EM Image Stacks with Learned Shape Features", IEEE Trans. Medical Imaging, vol. 30, no. 11, Feb 2011, pp. 474-486.

[8] Prasad, M. V. D., et al. "Fuzzy Classifier for Continuous Sign Language Recognition from Tracking and Shape Features." *Indian Journal of Science and Technology* 9.30 (2016).

[9] Kishore, P. V. V., et al. "Edge and texture preserving hybrid algorithm for denoising infield ultrasound medical images." *Journal of Theoretical and Applied Information Technology* 86.1 (2016): 120.

[10] Raghava Prasad, Ch, and P. V. V. Kishore. "Performance of active contour models in train rolling stock part segmentation on high-speed video data." *Cogent Engineering* (2017).

[11] Kishore, P. V. V., et al. "similarity assessment of 30 world sign languages and exploring scope for a sign – to – sign translator" *International journal of control theory and applications,* 10.11 (2017). pp. 315-335.