

A Review on Apple Detection Methods for Harvesting Robot

Kinjal V. Joshi

*G H Patel College of Engineering and Technology
Vallabh Vidyanagar, Gujarat, India
kinjalvjoshi@gmail.com*

Abstract

Nowadays robot is used to harvest fruit from trees. One major difficulty in developing system for selectively harvest fruits is to define the location, size and ripeness of individual fruits. These provisions are required to guide a mechanical arm towards the object. The computer vision approaches used to recognize a fruit depends on four basic features i.e. intensity, colour, shape and texture. In this paper, a review of different approaches to recognize and locate the apples is presented.

Keywords: *Apple detection, Colour based analysis, feature extraction, Machine learning*

1. Introduction

Currently agricultural automation has increased productivity by reducing manual labour and production costs, increasing yield and quality, and enabling better control over environmental effects. Even though these developments, various agricultural tasks are still being handled manually. Reducing agricultural labour energy which is physically hard, repetitive and time-consuming operations is still too challenging task for present-day robotic systems.

A serious feature in successful harvesting robots is their ability to process sensory information and to analyse and interpret visual input. The challenges like objects of various colours, shapes, sizes, textures, and reflectance properties; highly unstructured scenes with large degree of uncertainty; ever-changing illumination and shadow conditions; severe occlusions; and the sheer complexity of the typical unstructured agricultural scene, are the part of the problems that such a machine vision system must face. Because of these challenges, present-day success is still limited, and so agriculture is an important frontier of applied computer vision. Many applications have been developed for various tasks like autonomous navigation and obstacles avoidance, precision and selective spraying, weed detection, yield estimation, planting and ripeness and quality evaluation. The fruit detection can be considered the most prevalent application [4].

Interestingly, although much attention was given to the visual processing problems during the last 30 years, no selective harvesting robot has ever reached commercial maturity. While this unfortunate outcome cannot be blamed solely on the failure of machine vision to handle the challenges posed by the unconstrained and unstructured agricultural environment, it is undeniable that successful machine vision is critical for achieving high detection rates of ripe fruit in real time. The challenge of developing a cost-effective robotic system for fruit picking has been taken up by researchers at several places in the world [1–3].

In this paper literature related different approaches for apple detection is summarised. The rest of the paper is organized as follows. Different visual aspects that affect apple detection are described in Section 2. Different methods of apple detection with results are described in Section 3. Finally conclusion is represented in Section 4.

2. Visual Aspects

In this section different visual aspects that affect fruit detection methods are discussed. Different algorithms can be characterised by the visual aspects on which they operate [19].

2.1 Colour

Apples tend to have different colours than the leaves and branches around them so colour becomes one of the most popular visual factors used in harvesting robots that employ machine vision. But it can suffer from uncontrolled or changing illumination and shadows. Figure 1 shows effect of illumination on histogram of R, G and B component. Although the two apples hold the same perceptual colour, illumination variations even in the same scene greatly affect their measured colour components histograms *i.e.* R, G and B.

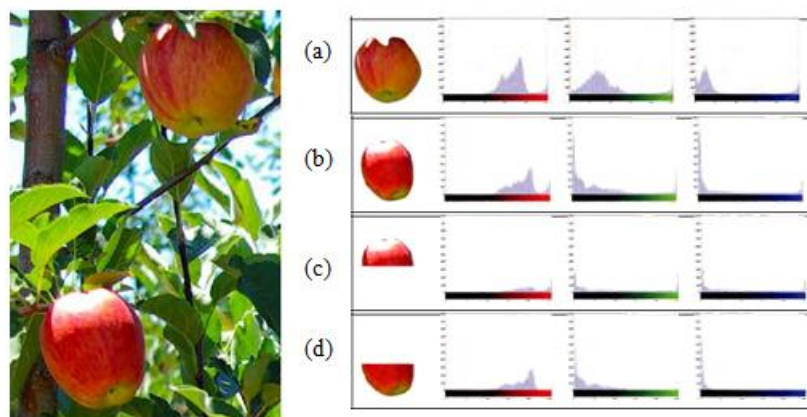


Figure 1. Histogram of R, G and B Component ([a] and [b] differences between different Apples in the Same Scene, and [c] and [d] differences within the Two Parts of the Lower Apple) [19]

2.2. Thermal Response

The other visual aspect can be considered is thermal response of objects, *i.e.*, their emitted radiation in the infrared range where it is strongly affected by both the temperature and the emissivity of materials.

Since leaves, unlike fruits, accumulate significantly less heat and emit it for a shorter time, thermography provides an excellent approach for target detection in harvesting machines [8]. However, since the thermal response is sensitive to the illumination and heat accumulation, fruit on different parts of the tree might respond differently [8-9]. Furthermore, detection problems may arise during specific parts of the day when fruit and leaf canopy have the same response. Naturally, issues related to the spatial organization of the visual data and the effects of occlusions are not addressed by this visual factor.

2.3. Texture

The texture property can be used as a discriminatory factor between different types of objects, *i.e.* between fruits and their surroundings [10-11]. The smooth skin of the fruit is utilized to distinguish fruit from foliage by using edge detection because fruit skin is smooth so less number of edges can be detected rather than leaf area. Hence, when the skin of the fruit is indeed smooth, patches of significant size and low edge density can represent the fruit quite strongly [10].

2.4 Shape

One global visual aspect that could be less vulnerable to illumination is the shape of the target, and although it is more computationally demanding to extract and analyze, shape is becoming increasingly more popular in harvesting robots. Since fruits are practically rigid objects, their shape relationship remains invariant in the three dimensional world. Although little invariance persists under perspective projection [12], shape remains a strong visual factor in the 2D image plane as well as under illumination variations.



Figure 2. Effect of Occlusion on Shape [19]

Despite robustness to illumination variations, shape is extremely sensitive to variations due to occlusions. Figure 2 shows that the shape of the projection of the apples changes drastically when occluded. Shape can be difficult to extract and analyze. It seems time consuming task also.

3. Methods of Apple Detection

In apple harvesting robot, the first key part is the machine vision system, which is used to recognize and locate the apples. It affects the robot's dependability and also determines its ability to directly, quickly and accurately recognize the apple in a real-time [4] [5]. In this section different methods for detecting apples from colour image are discussed.

3.1. Colour Based Analysis

The apples are determined with a certain threshold value as they are redish so having larger R component compared to G and B component. Most of these indicators are empirical parameters, and did not function well in some conditions. The indicator $2R-G-B$ is vulnerable to the impact of shadow, and the results are not ideal in cases such as those where back lighting is present or the background contains soil.

The other method based on colour is proposed by Si Yongsheng, Liu Gang and Feng Juan [5]. Here the most obvious feature *i.e.* red colour of mature apple is considered. To find a more effective colour indicator, line profiles of apple images in RGB colour space are plotted and analysed. Figure 3 represents line profile map. It shows that the pixels between M and N on the white line belonged to apple, and the other pixels belonged to the background, including leaves, stems and sky. Observing the profile curves, it is clear that in the apple, R values are greater than G values, while the opposite is true for the background. Attempting to use the R-G colour difference image segmentation method, the R-G difference is binary processed (if $R-G > 0$, $R-G = 255$, or $R-G = 0$), and the profile curve R-G is obtained. Apple can be recognized completely by R-G, while in the background area there are some

error recognitions. So G-B difference is also considered with R-G. If a pixel met $R-G > 0$ and $(R-G)/(G-B) > 1$, the pixel is identified as an apple, else the pixel is background.

This method is stable and less affected by light conditions. According to [5], total of 160 images captured under different lighting conditions were tested, and over 96% of the images were correctly segmented.

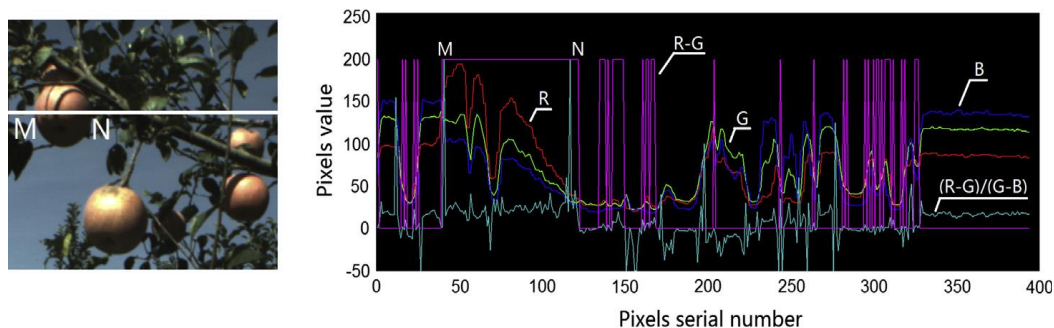


Figure 3. Result of Line Profile Map (a) Raw Image and (b) Line Profile Map [5]

3.2 Circle Detection Based on Random Ring Method (RRM)

This method is also proposed by Si Yongsheng, Liu Gang and Feng Juan [5].

Here as a feature the circle is detected because apples are circular in the images. Any three non-collinear pixels on the apple contour can determine a circle OA. Assuming that the center and radius of the circle OA are (O_x, O_y) and R_0 respectively. The radius of the apples in the images is within a certain range (30–60 pixels) due to the camera shooting distance limitations. The distance in this range is defined as the constraint radius. If R_0 exceed the above range, then it is discarded. Otherwise, the distances D_i between the points on the contour and the center (O_x, O_y) are calculated. Theoretically, these distances should be equal to the radius R_0 . The apple contours are generally irregular circles and pixels on the contour are not exactly on one circle. A width R_w is set and the number N_{max} of the contour points within $R_0 \pm R_w$ is counted. The higher the number N_{max} is considered as the better circle fitting. Circle fitting is performed several times. The circle corresponding to the largest number is selected as the final result.

This method is able to circle fitting even when the contours are not standard circle or when part of the contour is occluded. As for the apples appear in a cluster, there is possibility that the RRM will find a radius larger than 60 pixels through 3 random points. In that case, the RRM algorithm would discard it because that radius is out of range.

This method provides 89.5% accuracy on 160 tested images. The circle centers and radii are extracted precisely. Less than 20 mm error is reported with measuring distance between 400 mm to 1500mm

3.3 Method Based on Machine Learning

For machine learning two approaches Artificial Neural Network and Support Vector Machine are used for harvesting robot.

3.3.1. Artificial Neural Network

ANN has been proposed for classification between different elements in the scene such as fruit vs. background by Bulanon *et al.* [14]. An ANN is a formal computational model that is used frequently in machine learning. With combinations of all learning patterns like unsupervised, supervised, reinforcement, ANN has been quite popular in machine vision for harvesting robots. In such cases the input values (usually local colour features) are fed to the input layer of the network while the output layer provides a binary categorization. Training such networks is usually done in a supervised fashion by providing fruit and background examples [15] to build a look-up table that maps each point in the selected colour space to its corresponding class (fruit, background).

Wachs *et al.* [16] has proposed extravagant use of ANN. Here three ANN classifiers are used: One for each of the three colour spaces $L * a * b$, HSV, and RGB, in order to leverage the advantages that each one may provide. First, each of the three ANNs classifies the pixels of each image patch into several predefined classes (apple, leaves, branches, ground and sky). Then, a majority vote among the networks determines the final class that is associated with the patch.

3.3.2. Support Vector Machine

Rakun *et al.* [18] has used SVM to classify apples with prototype texture against other possible patterns. Apple detection is also done via LS-SVM, a version of SVM which solves a set of linear equations based on several colour and shape feature values [17].

3.4. Method Based on Intensity and Feature Extraction

This method is proposed by Sites and Dewilche [7]. This intensity-based method uses a B/W camera and colour filters (630 to 670 nm) to increase the contrast between the fruits and the background. Artificial light is used and most of the images are recorded under night operation. The whole method can be divided into five step: 1) setting threshold value for intensity, 2) smoothing by a binary filter, 3) segmentation by 8-neighbor connected component labeling, 4) feature extraction (area, perimeter, compactness, elongation), and finally 5) classification by a linear decision function or a nearest-neighbour method.

Classification results around 89%-90% are obtained for working at night and for mature fruits. During the day, 84% classification accuracy is declared and at least 20% of false detections.

3.5. Method Based on BP Neural Network and Hough Transform

This method is proposed by Xiao Changyi and ZhengLihua [13]. In this method first of all colour aspect is used. Here Neural Network modelling is selected to detect apple pixels and eliminate background pixels because colour distribution of apples and background are not completely linear. The considered other factor is shape. Because of overlapping and shading, the shape of an apple in the image may be shown as an incomplete circle. For example, if the apple is shaded by a tiny leaf, its shape maybe close to a circle. So, the circular curve is determined as the standard apple shape, and the Circle Hough Transform is proposed to detect the apples from tree images. In order to achieve detecting apples from tree images by using the Circle Hough Transform, edge detection is performed. The second step is to obtain the apple edge images which are the source materials for detecting apples. Figure 4 represents the detected edges.



Figure 4. Final Apple Edge Image [13]

Since the different apple surface textural properties, the aforementioned edge detecting algorithms has left small texture edge points inside the contour of apple, and accordingly those continuous pixel clusters less than 10 pixels in the edge image are removed. The third step is to implement Circle Hough Transform algorithm on each obtained edge image to detect apples as shown in Figure 5. Each circle and its center are detected through Circle Hough Transform algorithm, and the number of apples is then calculated by counting the centers of the apple circles. Figure 6 represents final apple detection.

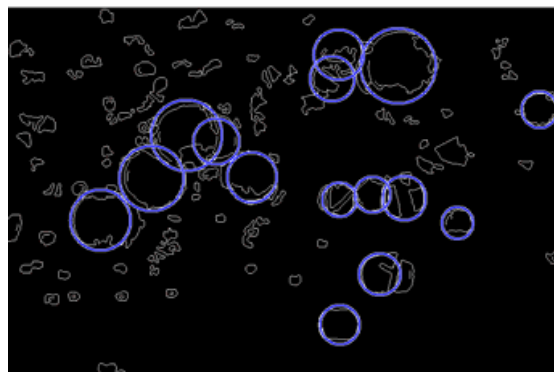


Figure 5. Hough Circle Detection [13]



Figure 6. Apple Detection [13]

This method has a strong robustness in dealing with the apple recognition problems such as illumination changes, shaded by leaves and branches, and overlaps with other apples. However, apple detecting algorithm would not work effectively if the area of shading or overlapping is too large.

The validation experimental results show that the correlation coefficient between this method and manual method is 0.985.

3.6. Method Based on Colour and Shape Feature Extraction

This method is proposed by Wei Ji *et al* [20]. It involves following steps.

3.6.1. Image Pre-processing

Because of natural environment, the original unprocessed colour apple image unavoidably includes noise that influences its quality. So in this method, first of all a vector median filter is applied to image enhancement as pre-processing. It can not only remove noise information effectively and highlights the apple fruit in foreground, but also maintains good image edges. The basic process of vector median filter includes three steps. First of all, achieve the average vector \bar{X}_{ij} of all known colour image pixel vector X_{ij} . Secondly, calculate the distance S_{ij} between \bar{X}_{ij} and X_{ij} . Finally, make the S_{min} , which is minimum of the S_{ij} , be the output value of the window central pixel. Here, by assuming that the size of original two-dimensional colour image setting is $M \times N$ pixel (M is the rows number, N is the columns number), a vector median filter with a $n \times n$ window is adopted to deal with the noise. The arithmetic is as follows.

a) Take the pixel as a vector $X(i,j)$. By letting $r(i,j), g(i,j), b(i,j)$ be three parameters standing for pixels in RGB colour images, then we have $X(i,j) = [r(i,j), g(i,j), b(i,j)]^T$ ($i=1, \dots, M, j=1, \dots, N$)

b) Calculate the averages of r , g and b , then we get the average vector of the window as

$$\bar{X}(i,j) = [\bar{r}(i,j), \bar{g}(i,j), \bar{b}(i,j)]^T$$

Where

$$\bar{r} = \sum_{i=1}^M \sum_{j=1}^N r(i,j) / n \times n$$

$$\bar{g} = \sum_{i=1}^M \sum_{j=1}^N g(i,j) / n \times n$$

$$\bar{b} = \sum_{i=1}^M \sum_{j=1}^N b(i,j) / n \times n$$

c) Calculate the distance S_{ij} between each vector and the average vector, and obtain the minimum distance S_{min} of them.

$$S_{ij} = \| X_{ij} - \bar{X}_{ij} \|$$

d) Make the pixel X_{min} corresponding to S_{min} as the vector median of the window, then let it replace the central pixel vector of the window.

Here the size of original colour image setting is 640×480 pixel. Vector median filter of colour images with the $n \times n$ window being 3×3 , is adopted to carry out enhancement for the apple colour image. The image filter result is shown in Figure 7.



(a) The Original Image

(b) Image after Filtering

Figure 7. Results of Vector Median Filter Applied on Apple Colour Images [20]

3.6.2. Image Segmentation

The segmentation effect of seeded region growing is determined by the selection of initial seed point and the growing rules. Here the image pixels are mapped into column diagram of RGB model. Then, some pixels are selected as the preparative growing seed in these colours whose probability is the biggest. After that, the 5×5 pixel region of the preparative seed is tentatively selected and threshold value is set. Then the number of pixels which are less than the threshold value is calculated. If the number of pixels is more than 20, the region around these pixels is considered as the goal object, namely the apple fruit. If the number is less than 20, it is considered that the preparative seed picks some isolated points whose colour is close to the apple. Then, this preparative seed is rejected and a new one is chosen. Finally, a growing seed is selected to grow under the growing rules. If the colour difference between growing point and that of the central point is less than the threshold value, then this pixel is added to the growing region. After that, the central point is labelled. Whether label or un label its neighbored pixels is according to their colour homogeneity. For a colour image, to judge whether the colours are homogeneous or not can be measured by the Euclidean distance. It means that there are two pixels $P1 (r1, g1, b1)$ and $P2 (r2, g2, b2)$ in RGB model, and their distance is

$$d(P1, P2) = \sqrt{(r1 - r2)^2 + (g1 - g2)^2 + (b1 - b2)^2}$$

Based on the method proposed above, the apple fruit image can be divided into two parts including fruit and background. The segmented image is used to extract features. However, during the process of segmenting, the isolated dots, burrs and holes usually exist in the image. To decrease the influence of these noises for subsequent image recognizing steps, opening and closing operations are used in mathematical morphology to remove the noises. The segmentation and mathematical morphology operation results of apple image are shown in Figure 8.



(a) The Original Image (b) Mathematical Morphology Operation Results

Figure 8. Segmentation and Mathematical Morphology Operation Results of Apple Fruit Images [20]

3.6.3. Image Feature Extraction

As image features colour and shape two features are extracted.

Colour Extraction

Here HIS model is used. Compared with the RGB model, the HIS model is easier for one perceiving the colours, and also accords with people's habits to describe colours. This model is based on two important facts: (a) the component of I does nothing with the colour information of image; (b) the chroma components of H and S have close relation with people manner to perceive colours. In view of the varieties of the light intensity, the influence of the light and shade in the image should be avoided. The components of H and S which do nothing with lightness are chosen to extract image features. Each point in the HIS colour space of apple fruit image can be regarded as a point in three-dimensional space of HIS model. Therefore, the difference of two colours can be measured by the Euclid distance of two colour points. The chromatic aberration formula is as follows

$$\Delta E = \sqrt{(\Delta H)^2 + (\Delta S)^2}$$

By this way, the colours can be compared directly in HIS colour space, and the colour and intensity information can be controlled well. Moreover, it is efficient in distinguishing small difference of the colours. Thus H and S components in HIS colour space are employed to recognize the colour feature.

Shape Extraction

In this method, the feature vectors are extracted that can satisfy RST (round, scale, transfer) invariability. In all the features of apple, feature parameters of round variance, ellipse variance, tightness, and ratio of perimeter and square area can describe furthest the outline feature of apple efficiently, so the above four feature parameter vectors are extracted. The shape is extracted by following three steps (1) Calculate image outline position (2) Calculate parameters of outline region area and perimeter and (3) Extract feature parameter vectors of the round variance, ellipse variance, tightness, ratio of perimeter and square.

3.6.4. Support Vector Machine

Here for automatic apple recognition SVM is used. 150 apple fruit images are used to train SVM. Results based on only colour features, only shape features and combined features of colour and shape are verified. Polynomial, sigmoid and RBF kernels are used.

Using combined features of colour and shape with RBF kernel give highest accuracy *i.e.*93.3%.

4. Conclusion

In this paper different visual aspects for apple detection are discussed. Using single factor or combination of different factors, detection of apples from colour image is possible. Here different methods for apple detection and their results are represented shortly. Still any system does not match the capability of the human labour faultlessly, especially in the vision and the recognition capability. The real time cost-effective and fully automatic robotic harvester might have a long way from the final prototype yet.

References

- [1] Sarig, "Robotics of fruit harvesting: a state-of-the-art review", J AgricEng Res; vol. 54, no. 2, (1993), pp. 265-80.
- [2] "Special Issue, Agricultural robotics", Autonomous Robots, vol. 13, no. 1, (2002).
- [3] G. Muscato, M. Prestifilippo, A. Nunzio and R. Ivan, "A prototype of an orange picking robot: past history, the new robot and experimental results", Industrial Robot"; vol. 32, no. 2, (2005), pp. 128-38.
- [4] D. M. Bulanon, T. Kataoka, Y. Ota and T. Hiroma, "A segmentation algorithm for the automatic recognition of Fuji apples at harvest", Biosystems Engineering, vol. 83, no. 4, (2002), pp. 405-41.
- [5] S. Yongsheng, L. Gang and F. Juan, "Location of apples in trees using stereoscopic vision", Computers and Electronics in Agriculture-Elsevier, vol. 112, (2015), pp. 68-74.
- [6] W. Ji, D. Zhao, F. Cheng, B. Xu, Y. Zhang and J. Wang, "Automatic recognition vision system guided for apple harvesting robot", Computers and Electrical Engineering-Elsevier, vol. 38, (2012), pp. 1186-1195.
- [7] Sites and Dewilche, "Computer vision to locate fruit on a tree", Transactions of the ASAE, vol. 31, no. 1, (1988), pp. 257-263.
- [8] D. M. Bulanon, T. Burks and V. Alchanatis, "Image fusion of visible and thermal images for fruit detection", Biosystems Engineering, vol. 103, vo. 1, (2009), pp. 12-22.
- [9] D. Stajanko, M. Lakota and M. Hocevar, "Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging", Computers and Electronics in Agriculture, vol. 42, no. 1, (2004), pp. 31-42.
- [10] H. Okamoto and W. Lee, "Machine vision for green citrus detection in tree images", Environment Control in Biology, vol. 48, no. 2, (2010), pp. 93-99.
- [11] J. Rakun, D. Stajanko and D. Zazula, "Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multiview geometry", Computers and Electronics in Agriculture, vol. 76, no. 1, (2011), pp. 80-88.
- [12] D. Forsyth and J. Ponce, "Computer Vision: A Modern Approach", Prentice Hall, Upper Saddle River, NJ, USA, (2002).
- [13] X. Changyi, Z. Lihua, L. Minzan, C. Yuan and M. Chunyan, "Apple detection from apple tree image based on BP neural network and Hough transform", International Journal of Agricultural and biological Engineering, vol. 8, no. 6, (2015), pp. 46-53.
- [14] D. M. Bulanon, T. Kataoka, H. Okamoto and S. Hata, "Determining the 3-d location of the apple fruit during harvest", in Proceedings of the Automation Technology for Off-Road Equipment Conference, (2004).
- [15] A. Plebe and G. Grasso, "Localization of spherical fruits for robotic harvesting", Machine Vision and Applications, vol. 13, no. 2, (2001), pp. 70-79.
- [16] J. Wachs, H. Stern, T. Burks and V. Alchanatis, "Low and high-level visual featurebased apple detection from multi-modal images", Precision Agriculture, vol. 11, no. 6, (2010), pp. 717-735.
- [17] D. Kong, D. Zhao, Y. Zhang, J. Wang and H. Zhang, "Research of apple harvesting robot based on least square support vector machine", in Proceedings of the International Conference on Electrical and Control Engineering, (2010), pp. 1590-1593.
- [18] J. Rakun, D. Stajanko and D. Zazula, "Detecting fruits in natural scenes by using spatial-frequency based texture analysis and multi-view geometry", Computers and Electronics in Agriculture, vol. 76, no. 1, (2011), pp. 80-88.
- [19] R. Mairon, "Computer vision for fruit harvesting robots-State of the art and challenges ahead", International Journal of Computational Vision and Robotics, (2012).

Author

Kinjal V. Joshi, is working as an Assistant Professor in Computer Engineering department of G H Patel College of Engineering & Technology, Vallabh Vidyanagar, Gujarat, India since 2007. Her research interests are in Image Processing, Artificial Intelligence and Computer Vision. She has published various research papers in these areas.

