International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Handwritten Telugu Character Recognition & Signature Verification

Mr.R.Venkatesh[1],B.Sri Harshitha[2],Ch.P.P.Priya Likhitha[3],G.Hemalatha[4],B.Uma Rajasekhar Guptha[5]

[1]Assistant Professor, Department of Computer Science
[2-5] B. Tech Student, Department of Computer Science
[1-5] Raghu Engineering College, Visakhapatnam

Abstract - Behaviour reputation stands as one of the earliest applications in sample reputation. While spotting handwritten characters is an clean venture for humans, it is a formidable task for computer structures. Optical Character Recognition (OCR) is an crucial answer primarily based on optical systems, which enables automatic reputation of scanned and digitized characters This paper explores into optical man or woman popularity strategies in particular developed for handwriting Telugu within the characters. Telugu, a Dravidian language spoken especially in Andhra Pradesh and Telangana, India, offers precise challenges because of its complex alphabet Basic parts of Telugu script together with "vattu" which stands for vowels and "gunitalu" which means that tone the complicated syllables add to the complexity. Combining OCR strategies with Harris corner popularity, the paper affords insights into the accuracy and efficiency of handwritten Telugu person reputation and the fidelity of handwriting This have a look at contributes to the development of character reputation in particular on in complex written languages which include Telugu and gives realistic explanations for handwriting verification processing.

Key Words: Optical Character Recognition (OCR), Telugu, Vattu, Gunitalu, Harris corner detection, Handwritten Character Recognition, Signature Verification.

1.INTRODUCTION

Telugu, a Dravidian language, is the official language of the Indian states of Telangana and Andhra Pradesh. It is one of the Republic of India's twenty-two official languages. Telugu is the second most spoken language in India and the fifteenth most spoken language globally. Given that Telugu is one of the world's oldest languages, the Indian government has designated it as a classical language. The Telugu script is also widely used for writing Sanskrit texts and bears significant similarities with the Kannada script, as both scripts share a common ancestry. However, because of its complexity, character recognition is highly challenging.

Among the first uses of pattern recognition is behaviour recognition. For humans, handwriting character recognition (HWC) is an easy assignment, but a difficult task for computers. When it comes to creating handwritten character recognition systems (HWCR) for Telugu text, not much significant work has been published previously. Research on character recognition is particularly popular since it has many potential applications in areas such as automated library construction, language processing, banks, post offices, security agencies, and multimedia production. It can also help the blind read.

These days, hand-drawn lines are incredibly common in graphic design, pattern recognition, and computer vision. Optical character recognition (OCR) automatically recognizes and digitizes scanned characters using an optical technique. The technology known as optical character recognition, or OCR, An electronic or mechanical gadget that sends pictures of written, handwritten, or printed text OCR, or optical character recognition, focuses on text, individual letters, or characters. The usage of paper documents, images, cellphones, iPads, and other devices for pattern recognition has expanded, making handwriting recognition (HR) an interesting area of research in recent years. The OCR system that has been put in place for English cannot be used for Indian characters since a single character in India can be a combination of vowels and consonants or a basic character made up of a single vowel or consonant. Therefore, compared to English, OCR is a very difficult task in the Indian setting. The development of OCR systems for Indian languages has been ongoing recently. When it comes to OCR, searches conducted in Indian languages lag far behind those conducted in any other language, including English, Chinese, and Japanese. There isn't much research on OCR, particularly for Telugu. Our goal is to create an OCR with Telugu compatibility and no limitations on input format.

A popular bio-metric asset is a handwritten signature. Unlike fingerprint technology, which can take time due to the user's ignorance, the primary benefit of manual handwriting over other identification technologies is that it can be used whenever the person is conscious and willing to write. Pattern recognition research is now being done in the domain of signature verification. Pen pressure, loop size, typing speed, typing speed and stroke, writing speed, pen pressure, loop size, and other handwriting features are used to identify a certain individual. feature extraction, verification, and signature recognition. By removing noise and other artifacts from the input images, the image preprocessing module is in charge of ensuring their quality.

Online signatures are typically preprocessed to remove changes that are thought to be superfluous to verification performance. Preprocessing processes including scaling, rotation normalization, and resampling are frequently performed. An enhancement procedure is used in the preprocessing stage to extract the signature's features. Before using any detection technique, various adjustments must be made to the image

signing. Various signature descriptions are used by signature verification systems. The traits that are chosen to be eliminated have a big influence on how accurate the signature verification procedure is. Because of the wide range of signature sizes and sample conditions, it is also the most difficult step in the signature verification process.

A significant bottleneck in any process involving the verification of signatures is the feature extraction procedure. Additionally, there is no assurance that two authentic signatures from the same person will be identical (intra-individual variation). The fact that proficient liars adhere to a true pattern—consistent variance amongst individuals—adds to the intricacy. This is not the case with fingerprints or irises, which differ greatly between two individuals. Interindividual variability ought to be far greater than intraindividual variability in an ideal world. Determining and eliminating the elements that raise interpersonal variability while decreasing interindividual variability is therefore crucial.

The test signature and the reference signature are compared using minimum similarity values, average of all similarities, and maximum of all similarities once the feature extraction method is applied. Choose any of the aforementioned similarity values to determine whether the signature is authentic or fraudulent. The reference and test signatures' threshold values are used for this comparison. It is considered to be a legitimate signature if the value is almost identical to the reference signal value; if the difference is greater than the threshold value, the signature is rejected. This cutoff point may be the same for every signature or it may vary for every one of them.

2. LITERATURE SURVEY

In their research, Velpuru et al. introduced an Optical Character Recognition (OCR) system tailored for Telugu text printed on paper. Velpuru et al. Telugu script written on paper with a suggested OCR. A grayscale image of the text is created during scanning. Average and average noise determination techniques are used for line noise segmentation. The zero-padding approach is employed for certain character sizes. State-of-the-art strategies like wavelet analysis and Hopfield networks were used to remedy Telugu individual reputation issues. There were several critical steps in the OCR technique.

The scanning method yielded grayscale images of the text. Line-noise segmentation changed into then achieved the use of each horizontal and vertical interpolation methods. A 0padding method turned into used to make certain uniform font length. Robust identity was performed via making use of wavelet analysis to extract heterogeneous visible records.

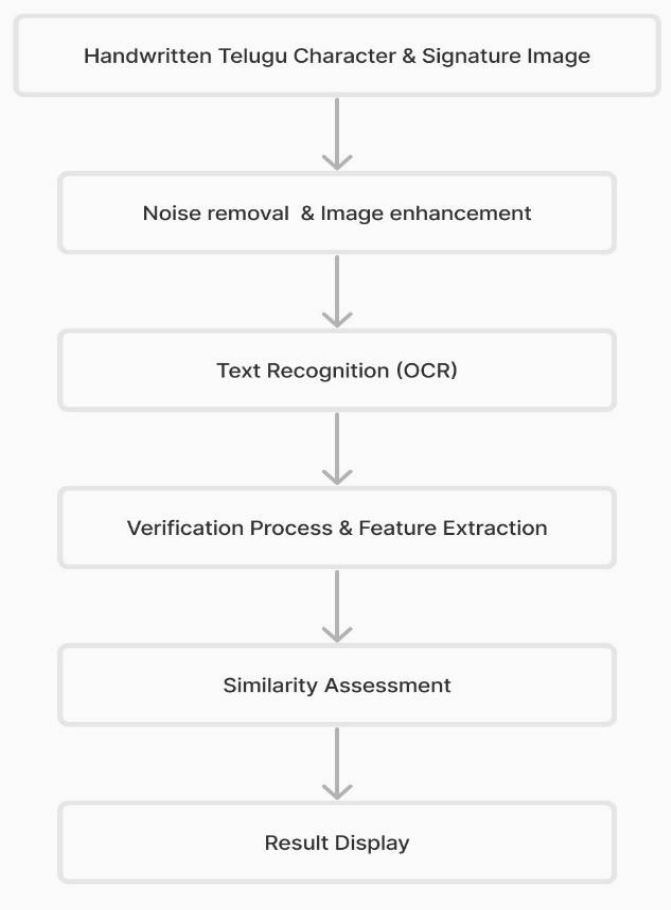
Using dimensional filtering, the photograph length turned into decreased to four 8 x 8 frames, which were then stacked to form frames. Signatures entered into enter symbols or sixty four bit textual content were the end result. Since every node inside the community become a Hopfield network, a dynamic signal may be detected. This technique didn't have too many problems with font-appearance modifications, while there have been a few icons that the gadget did not recognize efficaciously.

The Harris Research group on attitude detection-primarily based signature verification covers some of critical subjects. Signature verification is important in fields consisting of economic, legal, and identity verification, which require correct and reliable strategies. Among the face reputation methods, Harris face recognition has captured statistics approximately the capacity to correctly become aware of vital functions in images. Previous studies on signature verification has looked at a lot of strategies and techniques, with a massive recognition on improving device reliability and accuracy.

At this point, the introduction of the face is used in the handwriting verification functions, which shows the possibilities of choices, which have been applied to the influence of the prophets' construction and provisions methods of the advanced. Incidentally, the higher strength in accuracy uses the Harris corner personality to achieve.

3. PROPOSED SYSTEM

The figure below provides an overview of our suggested system.



The proposed system is aimed at development. A complete solution for Telugu writing Character recognition and signature verification. System architecture includes several key elements: Noise Removal and enhancement of images, text Recognition (OCR mechanism), verification Process, similarity evaluation and denote results.

Noise removal and image enhancement: Using image processing technology we tend to remove noise and improve input quality image containing handwritten Telugu characters

and signature. This step ensures: Follow-up recognition and verification process. It performs with clean, sharp images.

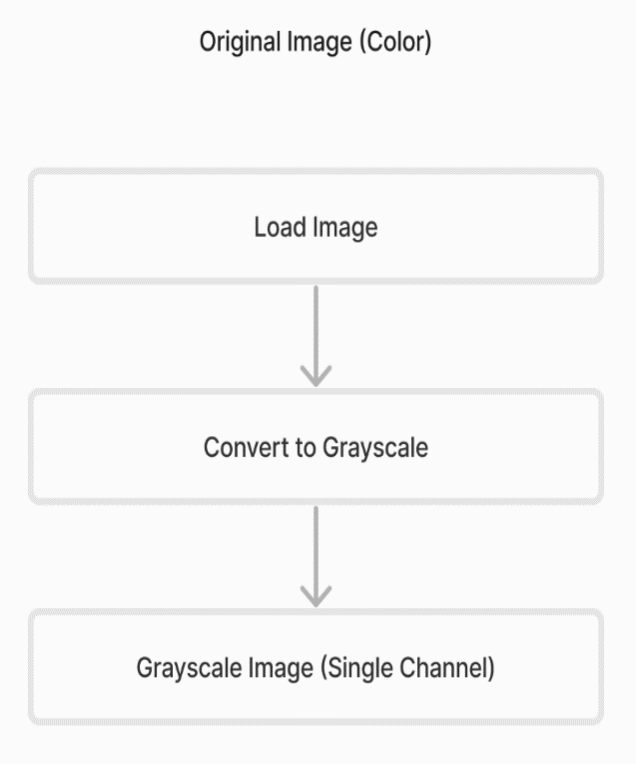
Text Recognition (OCR Engine): Identify and extract text areas preprocessed image using optical symbols recognition mechanism (OCR). Pixel analysis intensity distribution, shape and structure information about Telugu characters. Tesseract OCR The model serves as a standard for text recognition. Images provide a basis for comparison and improvement.

Review Process: Harris Implementation Corner Detection algorithm to extract the corresponding items function of input signature. Options such as: Speed, pressure, direction, spacing and size The signatures are extracted to characterize them. this step Provides reliable feature extraction for subsequent tasks.

Similarity evaluation: Because of the large character sets, the feature extraction technique is critical to the accuracy of recognition for optical technologies such as South Indian OCR. Comparison of extracted content Function between input signature and reference Signature using similarity score machine.Difference in indicators expressing the strength or quality of a characteristic measured accurately. If the difference in indicators is within the specified threshold, signatures are considered. Indicates agreement, acceptance.

Show Results: Displays test results indicates to the user whether the signature is: Accept or reject based on results Similarity evaluation. This is provided immediately. User feedback on authenticity the signature has been submitted.

For a diagrammatic representation of grayscale conversion of characters in picture format, we can illustrate the process step by step:



The diagram visually represents the transformation of a unique colour image into a grayscale picture, which is a critical step in person popularity and lots of other photograph processing applications.

Original Image (Color): This is the enter picture containing characters in image format. It consists of a couple of coloration channels (usually Red, Green, and Blue channels).

Load Image: The authentic image is loaded into reminiscence for in addition processing.

Convert to Grayscale: The loaded photo is transformed from its unique coloration layout to grayscale. This conversion involves reworking the photo right into a single channel representing the depth of every pixel, wherein darker pixels have decrease depth values and lighter pixels have higher depth values.

Grayscale Image (Single Channel): The ensuing grayscale photo carries best one channel representing the depth values of the pixels. This picture is suitable for diverse photograph processing obligations, which includes man or woman popularity, because it simplifies the processing whilst keeping critical capabilities.

4.IMPLEMENTATION MODULES

Module 1 : Character Recognition

Import libraries: OpenCV is a famous PC vision library used for the processing of images, including analysis, writing, and manipulation of images.

Pytesseract: For Google's Tesseract OCR engine, Pytesseract is a Python wrapper. This makes it easy to integrate Tesseract OCR into Python scripts.

Set Tesseract Path: Tesseract OCR calls for you to specify the route to the Tesseract executable document. It is necessary that pytesseract communicates with the TesserAct engine and performs an OCR operation.

Define function OCR: The 'recognize_characters' feature is designed to take advantage of the OCR characteristic. As an input, it's taking a photo channel. The software reads the image from a specified path, converting it to grayscale. By reducing the image to an uncoupled channel, grayscale adjustment makes it easier for your OCR technique and increases popularity of information. OCR is then executed at the grayscale photo the usage of Pytesseract. In order to ensure that text is as it should be extracted from Telugu pictures, the OCR device configures itself for picking up characters in Telugu. The program returns known colors that have been extracted from the image.

Specify image path: This step requires you to set a course for the photo report that must be used in conjunction with OCR. Ensure that the recommended course of action is correct and factor it into the preferred picture report.

Call OCR Function: Once the picture path is special, the recognize_characters function is referred to as with the photograph course as an issue. This causes an OCR process to

be triggered in which the feature scans a photograph, plays it and retrieves its text content.

Print results: Finally, the diagnosed characters obtained from an OCR method shall be revealed or displayed. This will allow you to see the text content extracted from your image and check its accuracy when performing an OCR operation. You can use the power of OpenCV and Tesseract OCR to extract text from photographs, especially if you specialize in detecting Telugu characters in this situation, by following these steps.

This may be useful for diverse programs inclusive of file evaluation, textual content extraction, and automation of statistics access obligations concerning Telugu language content.

Module 2 : Signature Verification

Read and show snap shots: Read grayscale photographs using an photo processing library such as OpenCV (cv2). Grayscale pics facilitate feature identification and matching with the aid of reducing computational complexity and specializing in intensity values rather than coloration. Specify grayscale photos to make them seen the use of capabilities like cv2.Imread() and cv2.Imshow().

Identify Harris capabilities: Use the Harris perspective detection set of rules to pick out feature points in grayscale photographs. Harris corner detection identifies corners or areas of interest in an photograph based on neighborhood intensity changes, figuring out critical structural features. Harris features in both images are detected with the usage of functions furnished by libraries which includes OpenCV on.

Filtering: A significant part of the overall signature identification process is feature extraction. To increase overall recognition rates, numerous feature extraction strategies have been put forth; nonetheless, the majority of them rely on the size and slope of the signature. Extract functions around regarded Harris corner points in each image.These features typically incorporate nearby picture bands or descriptors that document information approximately in the place of each function point.

Alignment: Alignment among photographs to align factors. Matching is normally achieved by evaluating function descriptors and finding the closest match using distance measures along with Euclidean distance or Hamming distance.

Get Matched Points: Retrieve the matched factors from the function matching technique. These matched factors represent corresponding features among the two pictures and shape the idea for assessing similarity.

Show Matched Features: Display the matched functions between the two snap shots to visualize the correspondences. This step facilitates in visually assessing the exceptional of feature matching and identifying any mismatches or outliers.

Check for Matching: Calculate the metric distinction for the matched factors to determine if they represent a valid fit. The metric distinction may be based on residences like distance, attitude, or depth between matched factors. If the absolute distinction falls beneath a sure threshold, don't forget the factors as matched; in any other case, they are considered no longer matched.

By following those steps, you may enforce Harris corner detection for feature matching between grayscale images and verify the similarity among them based totally on matched feature factors.

5.RESULTS

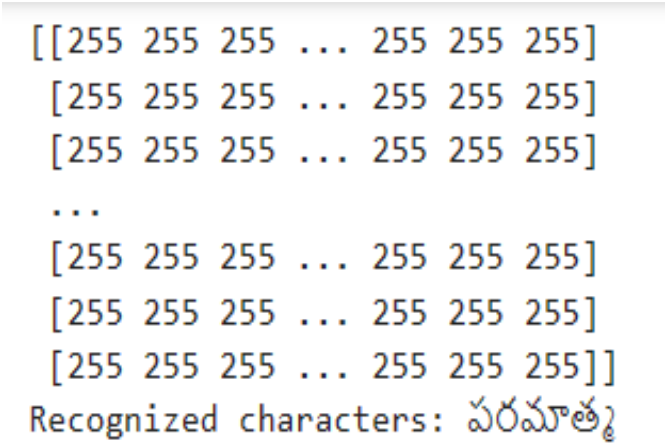


Fig-1: pixel values representation in the grayscale image and characters recognized from the Telugu word “paramaathma”.

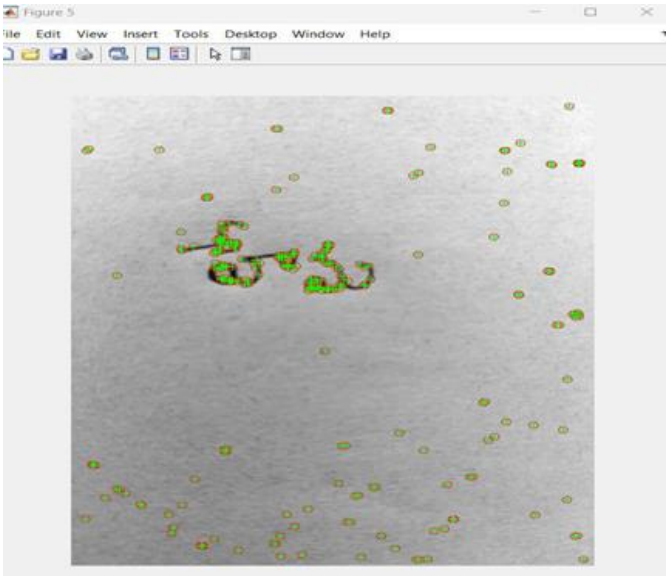


Fig-2.1: Overlapping plots of both genuine and test image.

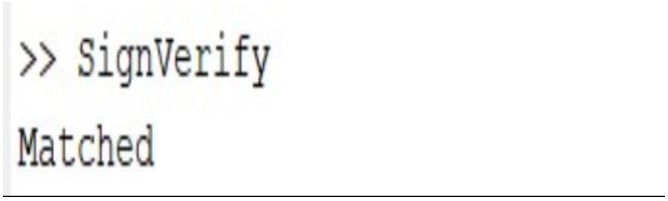


Fig-2.2: Output indicating that test signature is genuine.

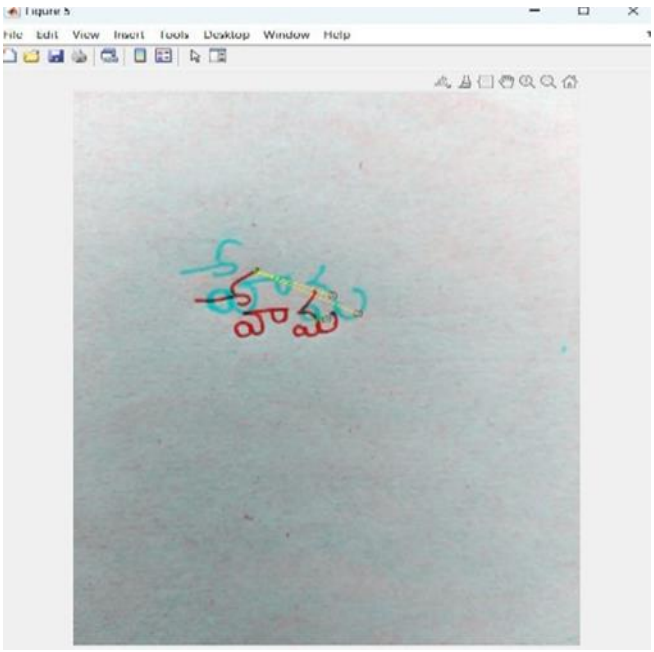


Fig-3.1: Difference in plots of genuine and test images

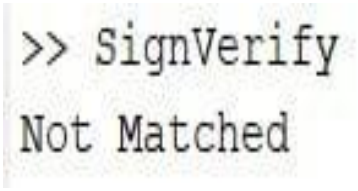


Fig-3.2: Output indicating that test signature is forged or fake.

6.CONCLUSION

End for Telugu handwritten characters validation and signature verification summarizes the findings and contributions The research, highlighting the importance of. The proposed system is what it means. In conclusion, this paper provides detailed information the way to deal with its challenges Handwritten Telugu characters found are Signature of authenticity by the combination of optical character and other advanced techniques Telugu character recognition compatibility (OCR) and feature extraction using Harris corner detection. We have made a solid schedule that we can. to recognize handwritten characters correctly .Check the accuracy of the manuscript holistically. We experimental results show that it is effective of the proposed system to obtain reliable results cross validation and verification performance various data sets. Also, about the system. The versatility and flexibility make it suitable for this applications that require authentication and literature review in Telugu context. Training is reduced Categorical cross-entropy and the SGD optimizer. It's astounding how much better the suggested method's results are than his viewpoint. The program in arts. The technique can then be expanded to numerous categories (such as gunitham and vattu) and can increase detection accuracy even in big data sets. Overall, this research It helps improve the paper using technology, in particular. Domain with handwritten Telugu script, open as such directions for future research in development recognition accuracy and expansion of the system internal capabilities to meet emerging challenges loyalty and loyalty activities. The training was carried out by decreasing the categorical cross entropy with the SGD optimizer.

7.REFERENCES:

[1]MR.V.CHANDRASEKHAR, B.HEMASRI "Handwritten Telugu Character Recognition Using Deep Cnn Model" Journal of Engineering Sciences Vol 14 Issue 08,2023.

[2]AISHWARYA SAHAI, AKASH KANT, “Online Signature Recognition And Verification In Telugu Script”, International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181,Vol. 9 Issue 10, October-2020.

[3]SRINIVASA RAO DHANIKONDA, PONNURU SOWJANYA, M.LAXMIDEVI RAMANAIAH, RAHUL JOSHI ,B. H. KRISHNA MOHAN, DHARMESH DHABLIYA, AND N. KANNAIYA RAJA , “An Efficient Deep Learning Model with Interrelated Tagging Prototype with Segmentation for Telugu Optical Character Recognition”, Hindawi Scientific Programming Volume 2022, Article ID 1059004.

[4]DR.ANUPAMA ANGADI, DR. VALLI KUMARI VATSAVAYI, DR. SATYA KEERTHI GORRIPATI, “A Deep Learning Approach to Recognize Handwritten Telugu Character Using Convolution Neural Networks”, Proceedings of 4th International Conference on Computers and Management (ICCM) 2018.

[5]“Telugu handwritten character recognition using zoning features” PANYAM NARAHARI SASTRY, TR VIJAYA LAKSHMI, NV KOTESWARA RAO, TV RAJINIKANTH, ABDUL WAHAB 2014 International conference on IT convergence and security (ICITCS), 1-4, 2014

[6]“Telugu handwritten character recognition using deep residual learning” BINDU MADHURI CHEEKATI, ROJE SPANDANA RAJETI 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 788-796, 2020

[7]H. LEI AND V. GOVINDARAJU, “A comparative study on the consistency of features in on-line signature verification,” Pattern Recogn. Lett., vol. 26, no. 15, pp. 2483–2489, Nov. 2005.

[8]“Online Handwritten Signature Verification System using Gaussian Mixture Model and Longest Common Sub-Sequences” SHASHIDHAR SANDA SRAVYA AMIRISETTI - Master Thesis Electrical Engineering September 2017.

[9] S. Z. LI, ENCYCLOPEDIA OF BIOMETRICS, 1st ed. Springer Publishing Company, Incorporated, 2009.

[10] A. K. JAIN, A. ROSS, AND S. PRABHAKAR, “An introduction to biometric recognition,” IEEE Trans. Cir. and Sys. for Video Technol., vol. 14, no. 1, pp. 4–20, Jan. 2004.