Yang Zhang and Hao Zhang HaoranLi [[1]](https://stacks.stanford.edu/file/druid:cg133bt2261/Zhang_Zhang_Li_Event_info_extraction_from_mobile_camera_images.pdf) described in their paper, an information extraction pipeline used for event flyers. **The major steps in the pipeline include image capture and upload, image preprocessing, text detection, OCR and NLP information extraction.** The paper lists several situations where a raw image could cause inaccurate results. **The OCR engine would assume the picture is taken from a perpendicular upright view, but images taken from a hand held camera could contain distortions.** **The illumination of the image not being uniform throughout and an image containing multiple blocks of text of different sizes and colours could also affect the output.** The image preprocessing methods included were edge detection, geometric correction(transformation), and Binarization.

***Rishabh Mittal and Anchal Garg [***[***2***](https://sci-hub.mksa.top/10.1109/ICIRCA48905.2020.9183326)***] introduced and explained the concept of OCR and the process of extraction by grouping it into majorly six steps: image acquisition, pre-processing, segmentation, feature extraction, classification, and post-processing. The paper reveals that the modern ocr system’s preprocessing pipeline is restricted to spatial image filtering, thresholding, noise-removal and skew detection/correction. Improving components like Scan goals, filtered picture quality, type of printer utilized whether ink-jet or laser, the nature of the paper, phonetic complexities, the lopsided brightening, and watermarks can impact the precision of OCR. Hence work can be done on improving the precision of OCR.***

***Sanjeev Kumar, Mahika Sharma, Kritika Handa, Rishika Jaiswal [***[***3***](https://www.ijitee.org/wp-content/uploads/papers/v9i7/G5745059720.pdf)***] proposed an algorithm to improve ocr accuracy with advanced image preprocessing using machine learning. Their focus was to reduce the noise of the image solely by scaling the original source image to around 300 DPI which has helped to eliminate the single biggest obstacle of the Tesseract, i.e Tesseract’s computation time of reading images with the highest character dimensions above 20 pixels. However, their algorithm does not cover images with uneven brightness, watermarks, or different fonts.***

***Dan Sporici, Elena Cușnir and Costin-Anton Boiangiu [***[***4***](https://www.researchgate.net/publication/341155551_Improving_the_Accuracy_of_Tesseract_40_OCR_Engine_Using_Convolution-Based_Preprocessing)***] underlined Tesseract 4.0 flaws, highly related to the segmentation procedure and proposed an adaptive image preprocessing step guided by a reinforcement learning model, which attempts to minimise the edit distance between the recognized text and the ground truth. This approach has boosted the character-level accuracy of Tesseract 4.0 from 0.134 to 0.616 and the F1 score from 0.163 to 0.729. The model adjusts samples with the purpose of maximizing the overall recognition efficiency without requiring external guidance or knowledge which has a direct benefit of including kernels, which can generate samples that might look unnatural. From a qualitative point of view, the changes are substantial yet not optimal since a reinforcement learning approach does not guarantee that local optimums will be avoided each time and hence the algorithm can get stuck on kernel configurations which will provide inferior results if not enough exploration is performed.***

***Sahana K Adyanthaya put forward a paper*** [***[5]***](https://www.ijert.org/research/text-recognition-from-images-a-study-IJERTCONV8IS13029.pdf) ***that presents the various steps taken to recognise text from images. The steps addressed in this paper were Image Preprocessing, Segmentation, Feature Extraction and Classification. The author highlights that the noise present in an image has a major role to play in successful text recognition and that noise removal increases the probability of accurate text recognition and generates more accurate output. The paper mentions that Gaussian filter and mean filter can be used for noise removal, that normalization should be done to ensure uniformity followed by binarization to convert the gray image into a binary image.***

***Naveen Sankaran and C.V Jawahar [***[***6***](https://projet.liris.cnrs.fr/imagine/pub/proceedings/ICPR-2012/media/files/2713.pdf)***] proposed a neural network-based framework that operates based on BLSTM-Bidirectional Long Short-Term Memory that allows OCR to work at the word level. It leads to over 20% better results when compared to a regular OCR framework. It uses a method that does not require segmentation, that is one amongst the foremost common reasons for the error. Also, it found an over 9% decrease in character error compared to the more widely available OCR framework.***

***Sai Abhishikth Ayyadevara, P N V Sai Ram Teja, Rajesh Kumar M [***[***7***](https://www.researchgate.net/publication/338569632_Handwritten_Character_Recognition_Using_Unique_Feature_Extraction_Technique)***],this paper deals with two different proposals of machine learning techniques. The first one was a new feature extraction technique, including the feature of three different existing feature extraction techniques. While the second one includes the analysis of the performance of three different neural networks for two different feature techniques- geometric and gradient. After doing all the survey, they concluded that the Convolutional neural network is most efficiently absorbed through the Levenberg-Marquardt algorithm.***

***Kukich[***[***8***](https://www.ijitee.org/wp-content/uploads/papers/v9i7/G5745059720.pdf)***] suggested using a n-gram dictionary or method based on the errors and returning the possible word to the dictionary using mathematical steps. These methods may reduce the total number of OCR errors in standard language names, but it is possible that the words may be correctly identified that are not in the dictionary of geographical names.***

***Dr. PL Chitra, and P Bhavani [***[***9***](https://poseidon01.ssrn.com/delivery.php?ID=543090115005102089070120100094082031000088051011052055098092024112029114031086103022038102053054023043125069124112070027120080105039082035013124027092026074000104116038055022079080008116116094069117005094015028065105070087120027122001104093019012068022&EXT=pdf&INDEX=TRUE)***], in this paper have studied various images to remove unwanted noise and performed enhancement techniques such as contrast limited adaptive histogram equalization, Laplacian and Harr filtering, unsharp masking, sharpening, high boost filtering and color models then the Clustering algorithms are useful for data logically and extract pattern-analysis, grouping, decision-making, and machine-learning techniques and Segment the regions using binary, K-means and OTSU segmentation algorithm. It classifies the images with the help of SVM and K-Nearest Neighbour(KNN) Classifier to produce good results for those images.***

Lavanya Bhaskar and R Ranjit [[10](https://sci-hub.se/10.1109/ICCCNT49239.2020.9225417)] discuss an event planner for the brochure images, that implements text extraction by convolution followed by MSER feature extraction and Stroke width method. The event planner then directly links the event text to the google calendar for scheduling the events. **However, the algorithm is not tested for event information taken from handwritten images and complex font text present in the images**

***Work by S. Akopyan, O.V. Belyaeva, T.P. Plechov and D.Y. Turdakov [***[***11***](https://www.ijert.org/research/text-recognition-from-images-a-study-IJERTCONV8IS13029.pdf)***] is based on a text extraction pipeline which is used to extract text from varied quality of images obtained from social media. Their work mainly focuses on dividing the input images into various classes and then preprocessing is done depending on the classes. This is followed by text recognition using the OCR engine. The dataset collected from social media is made use of in this work.***

***Anupriya Shrivastava, Amudha J.Deepa Gupta and Kshitij Sharma [***[***12***](https://www.ijert.org/research/text-recognition-from-images-a-study-IJERTCONV8IS13029.pdf)***] in their work have developed a system based on Convolutional Neural Network and Long ShortTerm Memory. The developed model identifies the texts from images which are horizontal, curved or oriented style. The model has four components. The first component performs feature extraction at the low level. The second component uses a shared convolution approach to extract high level features. Irrelevant features are ignored by the third component. The fourth component predicts the character sequences.***

Brijesh Kumar Y. Panchal, and Gaurang Chauhan [[13](https://iopscience.iop.org/article/10.1088/1742-6596/1973/1/012008/pdf)] proposed an implementation on the Android Application to extract using Tesseract OCR in which the following concepts are used, which are Adaptive Thresholding, Connected Component, Fine Lines, and Recognize Word. Using this Optical Character Recognition (OCR) Technology. The Application generates text which is printed on a clean, B/W or colourful background and then can be converted into a computer readable form ASCII. **With the help of this Android Application using Tesseract OCR, the system has two ways for Text Extraction. The first one is to capture a photo while the second one uploads an image from the gallery. After that the system can proceed as per the user requirement which portion of the image they want to crop or edit. After editing the picture, it converts into the text. This Android Application is for two languages, English and Hindi.**

K.Gaurav, Bhatia P. K. [[14](https://www.ijsdr.org/papers/IJSDR2009078.pdf)] , this paper deals with assorted pre-processing techniques used for handwritten recognition which consists of different images starting from a simple handwritten document and extending its radius to complex background and diverse image intensities. **The pre-processing techniques that were included are contrast stretching, noise removal techniques, normalization and segmentation, binarization, morphological processing techniques.** They came to the conclusion that no technique for preprocessing can single handedly be used to produce an image. **All the techniques go hand in hand. Even though after applying all the said techniques, the accuracy of the image is not upto the mark.**

Salvador España-Boquera, Maria J. C. B., Jorge G. M. and Francisco Z. M. [[15](https://www.ijsdr.org/papers/IJSDR2009078.pdf)], this paper outlines the hybrid Hidden Markov Model (HMM) is used to conceive the unconstrained offline handwritten texts. The main characteristics of the recognition systems is to produce a new way in the form of preprocessing and recognition which are both based on ANNs. **The preprocessing is used to clean the images and to enhance the non-uniform slant and slope correction.** **Whereas the recognition is used to estimate the emission probabilities.**

***K. Karthick, K.B. Ravindrakumar, R. Francis and S.Ilankannan [***[***16***](https://www.ijert.org/research/text-recognition-from-images-a-study-IJERTCONV8IS13029.pdf)***] have discussed the various steps in text detection in detail highlighting the different techniques used for the same. They have also emphasized on handwritten text recognition which is one of the complex fields. From their study it has been found that best results can be had with reduced computation time and it is possible to segment multilingual characters and enhance the character recognition rate.***