# COMIC POLYGLOT

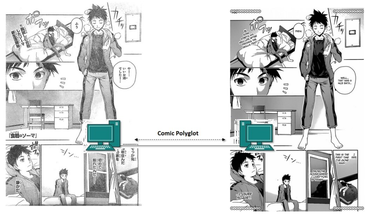
.Akshay . Priyanka . Gaurav . Harshvardhan . Aman . Farhat.

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA, SURATHKAL

Carnegie Mellon University

## Objective

Everyone loves comics. Wouldn't it be wonderful if we could read all the amazing comics spread across different languages like Japanese, Spanish, French and German in English or Hindi? Periodic comics in different languages come out every other day or week, but are inaccessible for a long time due to the intensive manual translations efforts required. This manual process consists of teams of people working separately on text extraction, translation and infusion for every line of text in a comic. Our project attempts to automate this process and make comics available in other languages within a day of original raw release with minimal human effort required. The hurdles that we face include automation of complex processes like text detection, extraction, removal, optical character recognition, language translation based on context and text infusion. This project on completion has the potential to revolutionize the scanlation process and make a difference in the comic industry.




## References

-Robust text detection in natural images with edge-enhanced Maximally Stable Extremal Regions

-Chen, H. Tsai, S.S. ; Schroth, G. ; Chen, D.M. ; Grzeszczuk, R. ; Girod, B

-LIBLINEAR: A Library for Large Linear Classification

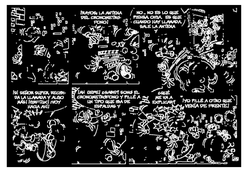
Rong-En Fan ,Kai-Wei Chang, Cho-Jui Hsieh,Xiang-Rui Wang,Chih-Jen Lin.

-An Overview of the Tesseract OCR Engine by R. Smith

## Text Extraction Pipeline

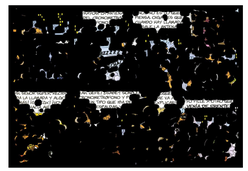
Original Image

Edge Enhanced MSER



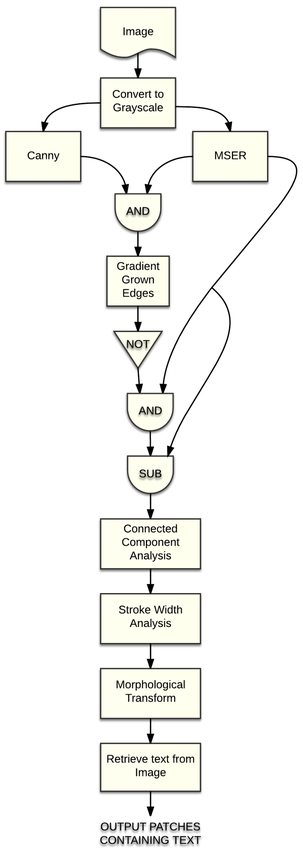
Gradient grown Edges

Connected Components Extraction



Stroke Width transform

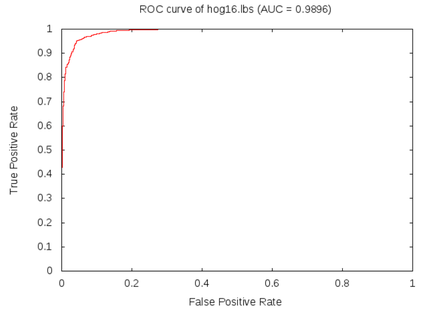
Output Patches




## SVM Classification

Images from the text extraction pipeline are separated into text and non-text patches via a Linear SVM classifier that has been trained on HOG features of the image patches with varying sizes of kernels( 2x2, 16x16 and 32x32 ).

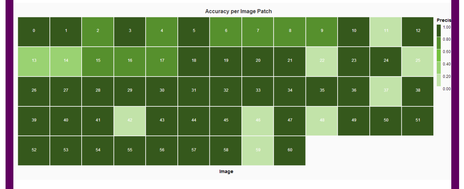
| | 2 x 2 HOG Kernel | 16 x 16 HOG Kernel | 32 x 32 HOG Kernel |
|---|---|---|---|
| AUC | 0.99925 | 0.98963 | 0.94925 |
| Recall | 95.21% (219/230) | 96.95% (223/230) | 89.13% (205/230) |
| Precision | 97.76% (219/224) | 98.23% (223/227) | 98.55% (205/208) |
| Accuracy | 98.49% ( (1047/1063 ) | 98.96% (1052/1063) | 97.36% (1035/1063) |




## OCR and Translation

Tool : Tesseract

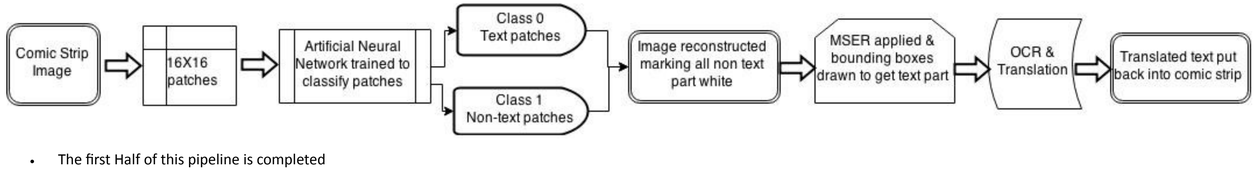
Cleaning using Dictionary of 90000 words by calculating Levenshtein Distance.




Translation tool

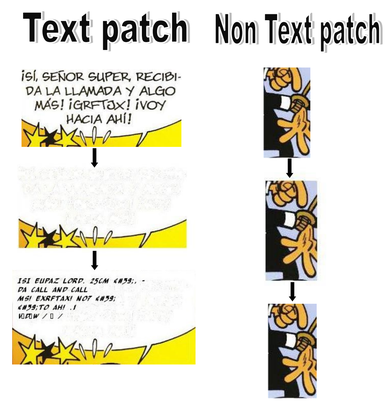
Google Translate API

## Neural Network Approach(Future work)



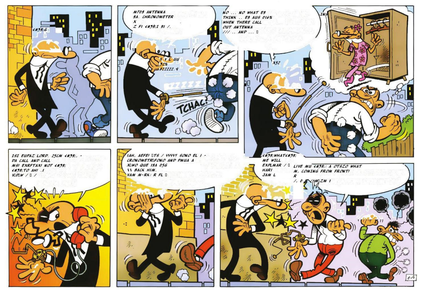

- The first Half of this pipeline is completed
- The rest will be completed in the future.

## Infusion




For every candidate patch, we remove the text regions if the OCR output is empty or the SVM classifies the patch as a non-text patch, else we leave it unchanged.

## Result



The current system is able to detect text in comics with reasonable accuracy of 83.87% (tested on 230 comic images). The OCR and translation needs to be improved a lot more to be able to give readable and understandable translations.

Under the guidance of

- Prof. Biksha Raj
- Prof. Rita Singh
- Pulkit Agrawal