1. Histogram: to enhance the contrast, simple histogram equalization is not good enough, because there maybe too many “white” region in the paper => would span the white region wider, and influence the result of the threshold. => we have to use the contrast-limited histogram equalization.
2. Font detection: if we reproduce a image with different font, then try to find the possible feature, is that possible??
3. Sharp the characters: maybe use the filter? High frequency filter to boost the high frequency??