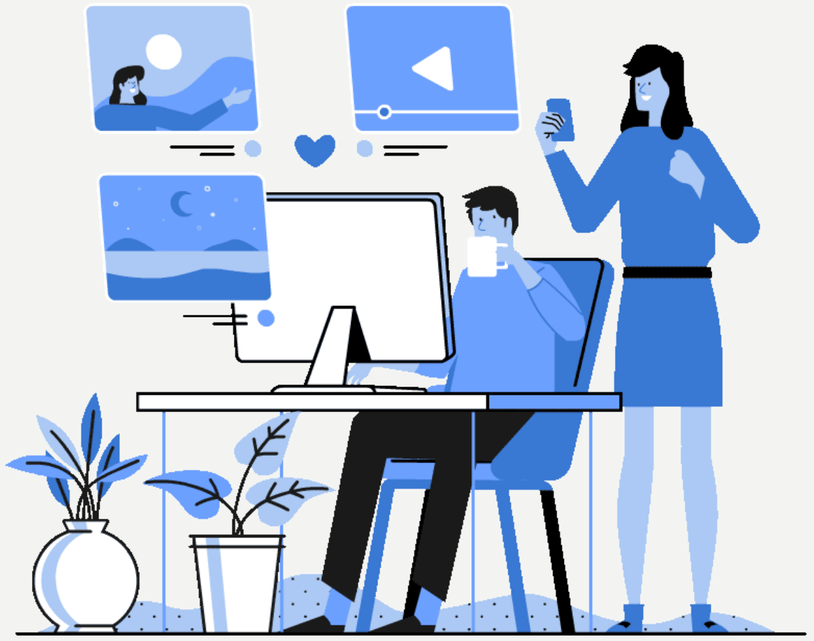
Using ML To Improve Image Quality

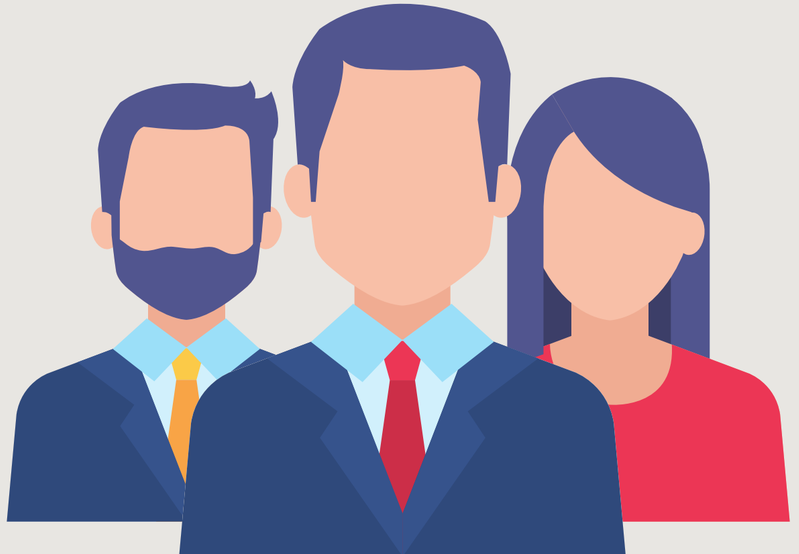
TEAM - 3

Table of contents



INTRODUCTION

Document scans don't always come out the way we expect. A common problem encountered when scanning documents is ‘noise’ which can occur in an image because of paper quality, the typing machine used, or it can be created by scanners during the scanning process.



Overview

When scanning pages, we sometimes scan the photos at odd angles, and such images are frequently found to be rotated and may lead to loss of texts.

When the scanning screen has not been adequately wiped, they can be noisy and full of speckle dirt markings.

We ensured that such flaws are eliminated and that the photos are readable by the reader in order to improve image quality.

To avoid speckle dirt stains across the images, it would have been more beneficial if the scanner had used optical character recognition (OCR).

What Is Image Noise?

Noise is an additional pixel values added to an image causing the loss of information.

Impulse Noise

Pixel values are completely different from the surrounding pixel values

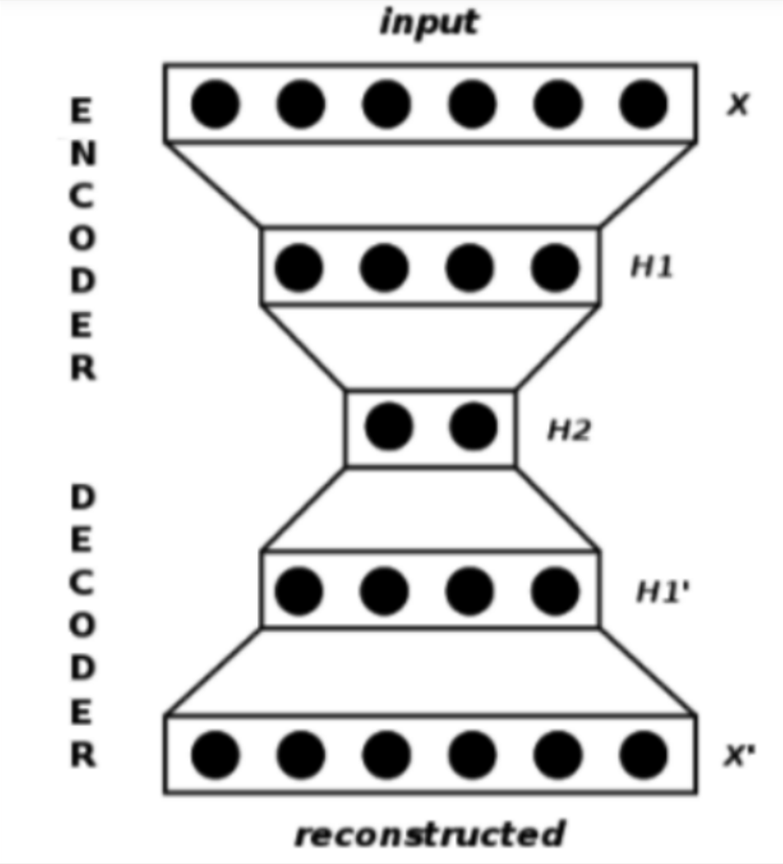
IMAGE NOISE

Additive White Gaussian Noise

Each pixel in the image will be changed from its original value by a small amount.

Image Denoising Using Autoencoders

Autoencoders are a class of Neural Networks that attempt to recreate their target using Back Propagation



ML Techniques

Number of different ways we can solve this problem:

- Denoising [In R we can use EBImage and Magick packages for denoising.]
- Blurring the image to some extent can help us to remove some noise, but it will also remove some important information.
- Gaussian Threshold method using openCV [Python]

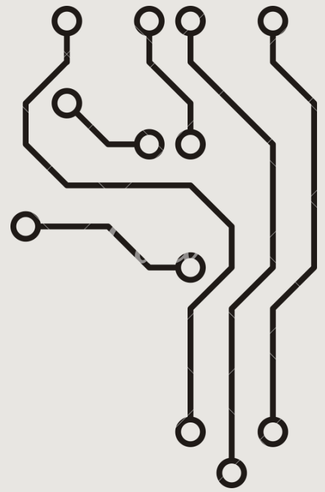
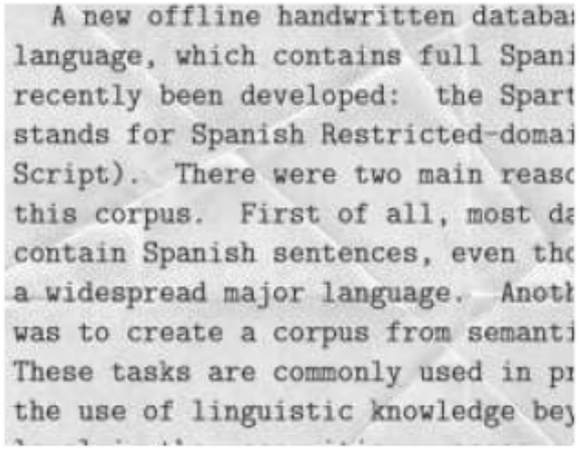


Image 1

Original Image 136.png



Processed Image Manually 136.png

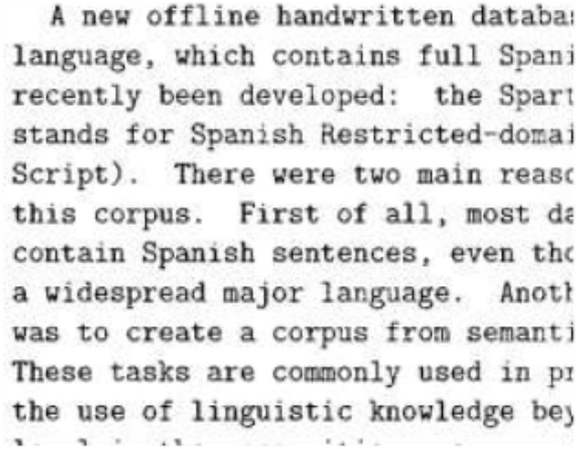
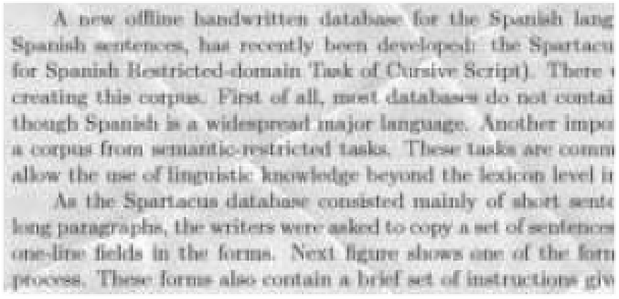
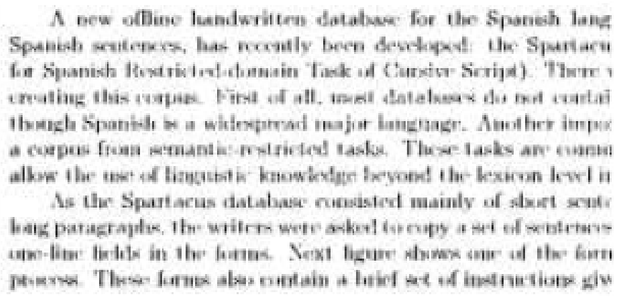


Image 2

Original Image 16.png



Processed Image Manually 16.png



IMPROVED IMAGES

After removing noisiness from the scanned images.


Thank you!