



INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY

H Y D E R A B A D

Sheet Music Reader

Digital Image Processing (CS7.404) - Project Report

Submitted By

Srikant Konduri – 2022201017

Subhadeep Biswas – 2022201039

Sharanya Bhattacharya – 2022201040

Hrithik Sagar – 2023900021

Task: We have, given a sheet music, tried to detect various types of notes, clefs and lines. We have diverged quite a bit from our reference paper, by Sevy Verma and Prateek Harris of Department of Electrical Engineering, Stanford University.

Methodology:

Segmentation and Preprocessing: We have used different Segmentation and Preprocessing steps as given below (applied in order):

1. Erosion using horizontal line of $1/5$ of the image width.
2. Row sum has been done to get a 1D vector.
3. Eroded with vertical lines of variable length.
4. Dividers are calculated by taking the middle of the black portions,
5. Staff lines are identified by thresholding the row sums. Before that opening and closing are used to make the staff lines smooth
6. Staff lines are removed for further processing.

We have used the following Sheet Music (without noise) for all our tasks:

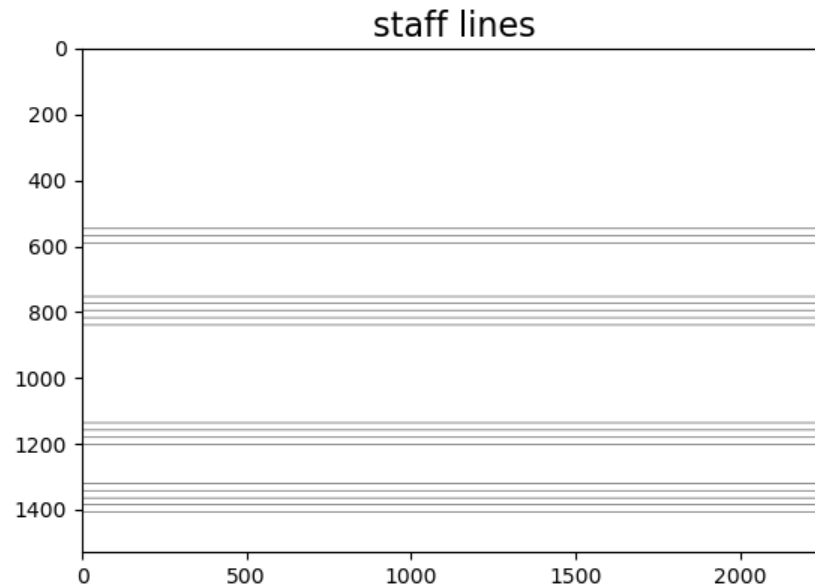
Free sheet music - Capotasto Music

Mary Had A Little Lamb

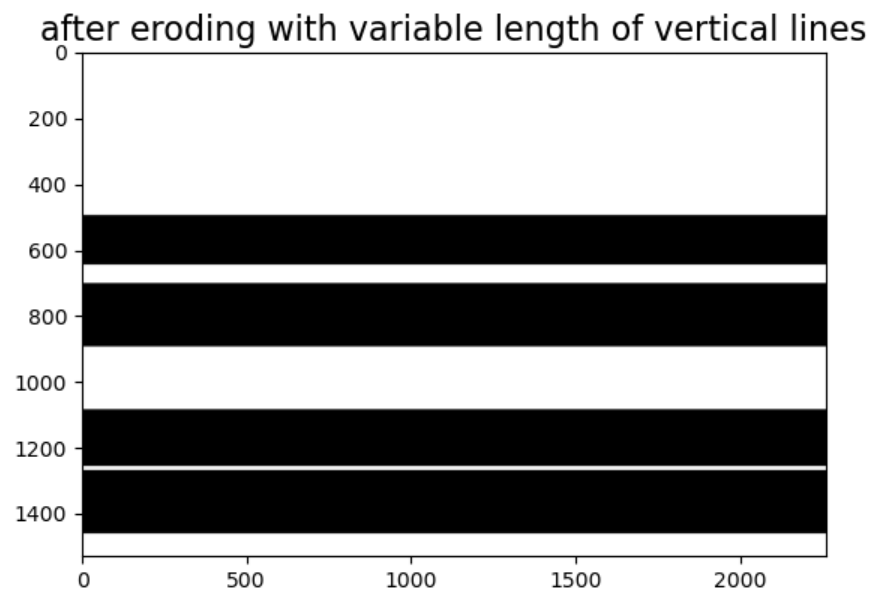
Lowell Mason



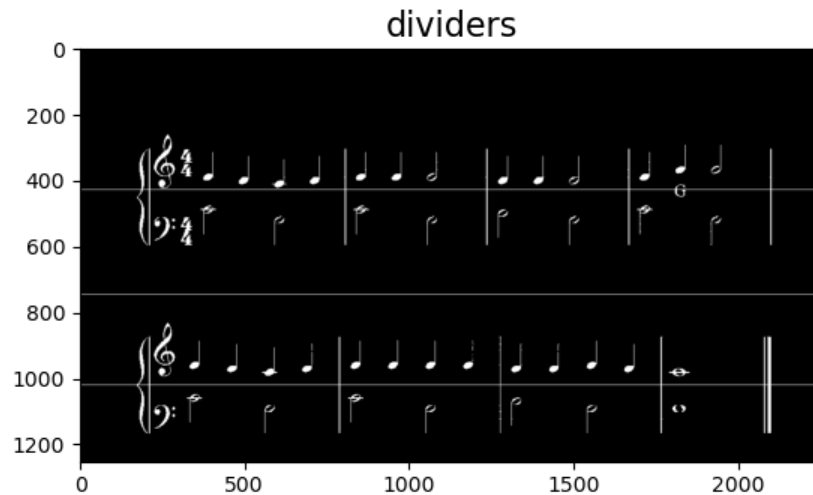
Results of Segmentations and Preprocessing: On performing the above steps we have received the following results:
After Step 1 and 2:



After Step 3:



After Step 4, 5 and 6:

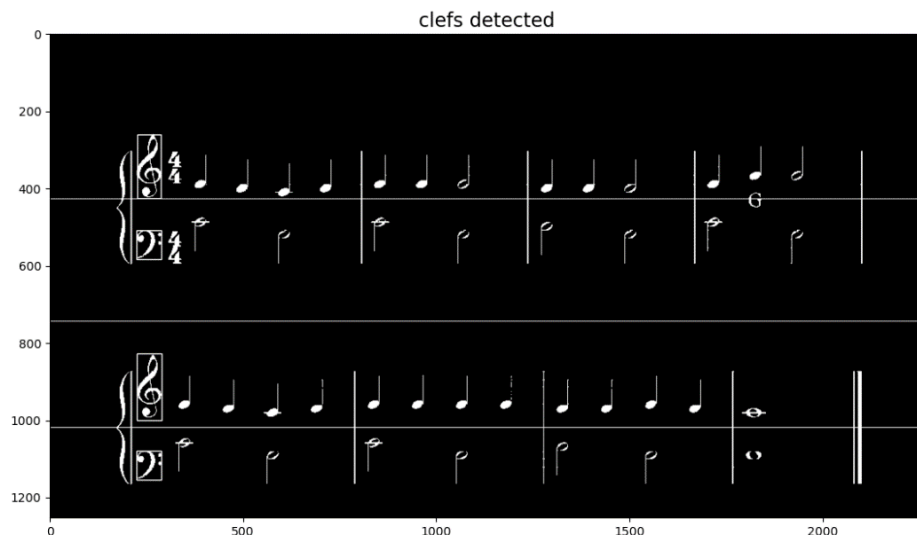


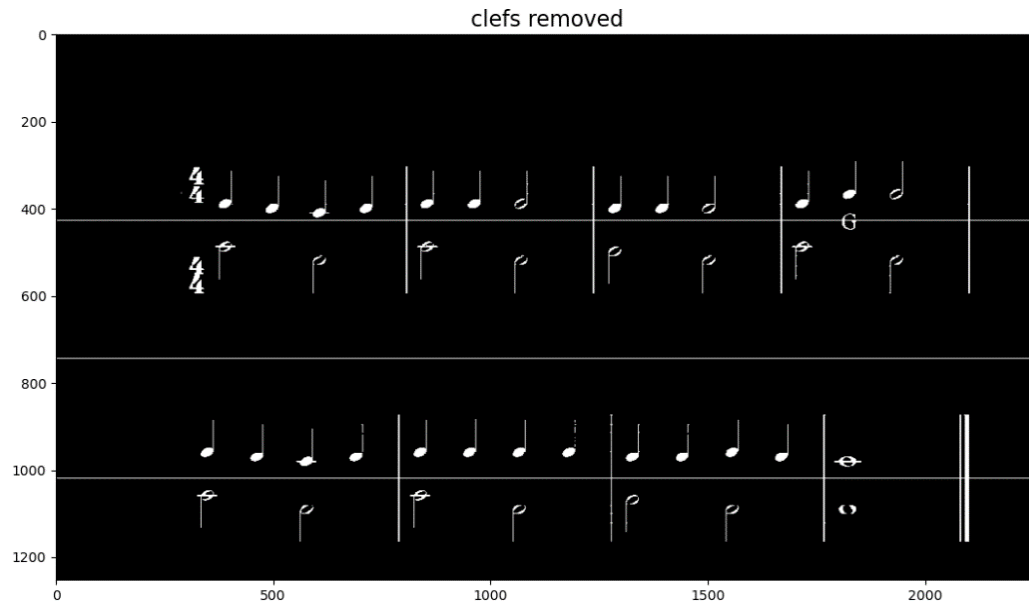
Clef Detection:

Method in Paper – Assuming Clef is always the first object drawn on the left side of the staff, Sum of the vertical columns is used to detect clefs.

Our Method – For most sheet music, the above assumption does not hold. Most of the sheet music use the same font for the clefs, so we have used template matching to detect the clefs.

Results of Clef Detection:





Notes Detection:

Methods mentioned in the Paper:

Quarter and Eighth Note Detection: Quarter and 8th notes are grouped together as both have filled black circles. Process = erode + dilate using disk SE (used for normalizing space between staff lines). Then, these results are thresholded by eccentricity and area. If variance is large enough, components more than 1.75 SD, then below median area are removed.

Whole note and Half Note Detection: Very similar to detecting quarter and 8th note method. Key Idea is Original Image - (fill up holes in an image). Result = image with non-filled notes present. Used Watershed algorithm. Drawback -> staff lines form closed regions without removal, which results in filling up entire regions hence we use a preprocessing step to remove staff lines, but this again opens up the connected regions which are difficult to fill and detect. So to avoid this, we take SE = proportional to the fixed ratio of the size of the width of the staff. Dilate it with the SE, gaps due to removal of staff lines are removed, original size of image is increased. Then, fill up closed regions using watershed algo. Now we subtract all these from original image

Our method:

Same as clef detection, we have used template matching to detect the quarter notes and half notes.

Results of Notes Detection:

Red Points are the detected Quarter Notes and the Blue points are the detected Half Notes.

