Project Report

USA
Device Name: ECG Machine

REF NML903060   Qty: 9

LOT: 34690

Introduction

The objective is to create an excel file with appropriate columns of Device Name, REF, LOT, Qty and symbols. EAch symbol in the image has a specific code. The symbols and the texts should be detected and entered accordingly.
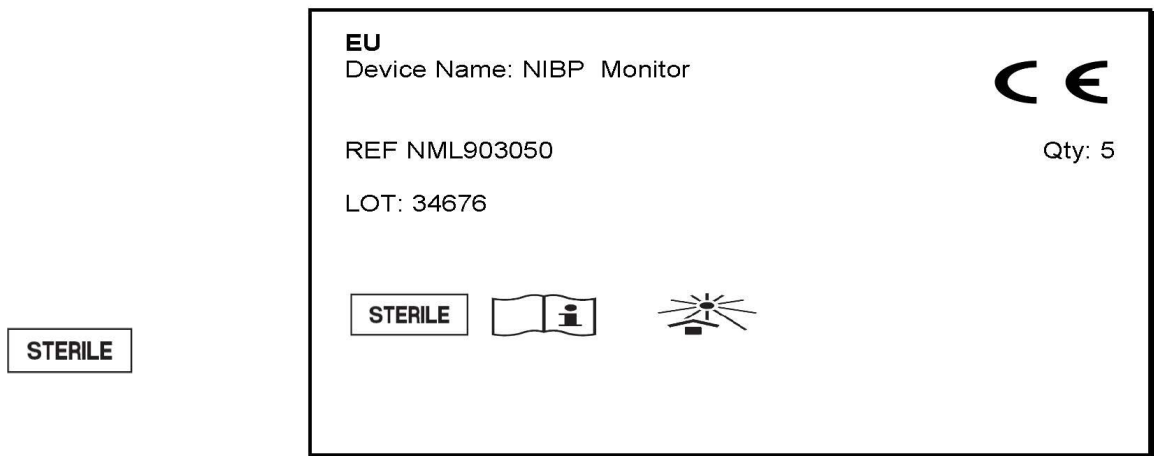
Approach

I first cropped the image into two parts using PIL. The first one contained the parts of device name, lot and ref. The other one with the quantity.

Qty: 9

USA
Device Name: ECG Machine

REF NML903060

LOT: 34690

I then used pytesseract to extract the text from the images and converted the string to a list to get the required words alone.

For the symbols part, I used matchtemplate from opencv to detect the symbols with the symbol as template and the result page as image. I appended the outputs accordingly with the appropriate code of the symbols. I looped through all the images by going through the images folder and looped again through all the symbols for each image to get the desired output. Since symbol 4 always appears in the top, and we

were asked to put the order of symbols from top to bottom, I iterated through symbol 4 first.

Template   **Image**

Output

Finally, after getting the appropriate texts and symbol codes for each image, I created a dataframe to add them all in the desired output format. I then created a csv file out of it to give the final output.

	A	B	C	D	E	F	G
1		Device Name	REF	LOT	Qty	Symbols	
2	0	Pulse Oximeter	NML903055	34683	4	2369	
3	1	Blood Warmer	NML903090	34641	1	1	
4	2	C-Pap Machine	NML903105	34662	1	15789	
5	3	ECG Machine	NML903060	34690	9	2578	
6	4	HFNC Machine	NML903095	34648	5	1289	
7	5	Infusion Pump	NML903065	34697	10	2589	
8	6	NIBP Monitor	NML903050	34676	5	417	
9							

Output

~~~ Thank You ~~~