Specific form elements recognizing:

Now, only for finishing now

Buttons

how is it done:

Tesseract has problem with negative parts, it’s color blind, that is useful for buttons recognized

select boxes

this is done by “hack” solo text recognized as “V” or “vV”

----

In the process - this three steps

* boxes ( also empty boxes ) recognizing by colors - histogram

tesseract don't recognize color areas

vectorizers / libraries - are not useful for this

I have my own version of “poor man puzzle” solution for detecting boxes

90% done

* tesseract “special fonts” - for checkboxes, radio buttons and states

this could be “killer feature” for using tesseract for this task

60% done

* combine all results together (intelligent selection and combining results for specific layers)

50% done