Tesseract OCR Installation:

## Prerequisites:

# MacPorts Installation:

MacPorts is an alternative to brew for installing OSX programs. The easiest way to install MacPorts on a Mac OS X system is by downloading the .pkg for Yosemite and running the system's Installer by double-clicking on the .pkg contained therein, following the on-screen instructions until completion. The installer can be downloaded from here:

<https://distfiles.macports.org/MacPorts/MacPorts-2.3.3-10.10-Yosemite.pkg>

The MacPorts “selfupdate” command will also be run for you by the installer to ensure you have our latest available release and the latest revisions to the “Portfiles” that contain the instructions employed in the building and installation of ports. After installation is done, it is recommended that you run this step manually on a regular basis to keep your MacPorts system always current:

* sudo port -v selfupdate

At this point you should be ready to enjoy MacPorts!

# Tesseract OCR Installation:

The easiest way to install Tesseract is with MacPorts. Once it is installed, you can install Tesseract by running the command

* sudo port install tesseract

Languages can be added with the command ‘sudo port install tesseract-<langcode>’. A list of available language codes can be found on MacPorts tesseract page. For example to install the English language we would do:

* sudo port install tesseract-eng

Once installed the command below will display the versions of tesseract and leptonica installed:

* tesseract –v

To process an image for OCR with tesseract the syntax is:

* tesseract inputfile.png outputFile

This will process the file ‘inputfile.png’ for OCR and save the output in a file called ‘outputFile.txt’.