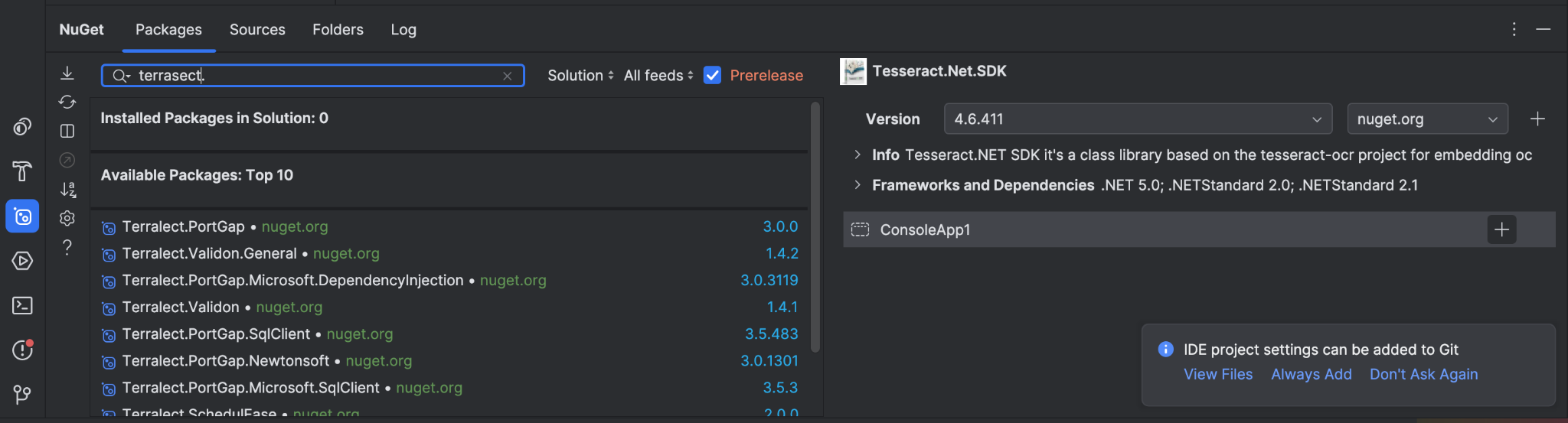
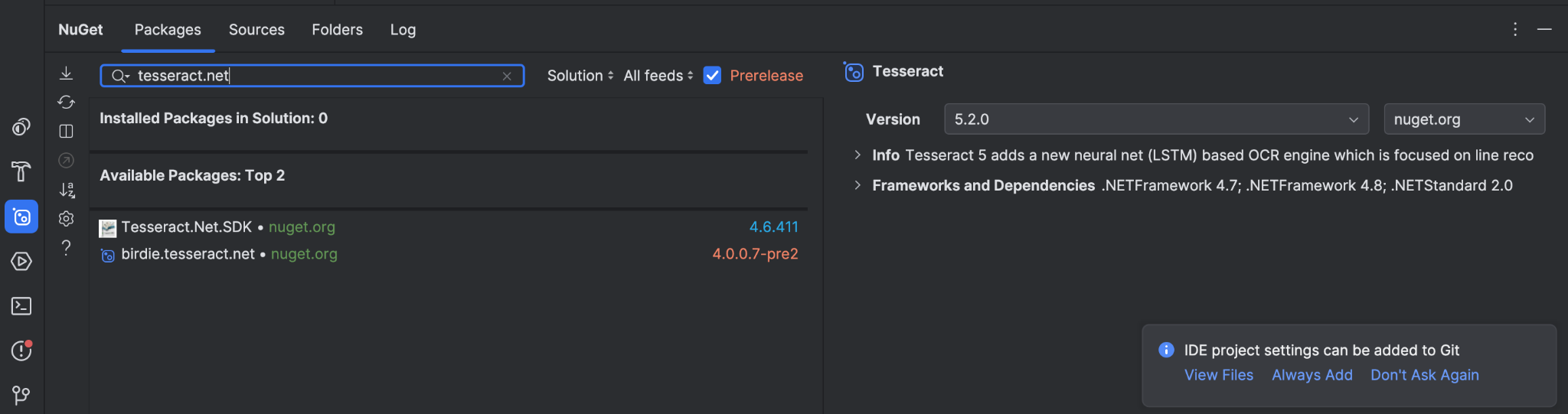


For our Project topic Creating Text from images with OCR API, it has been mentioned that we should use Terrasect. But upon research we saw that Tesseract is free and terrasect is paid and professor Dobric has told to only use free version.





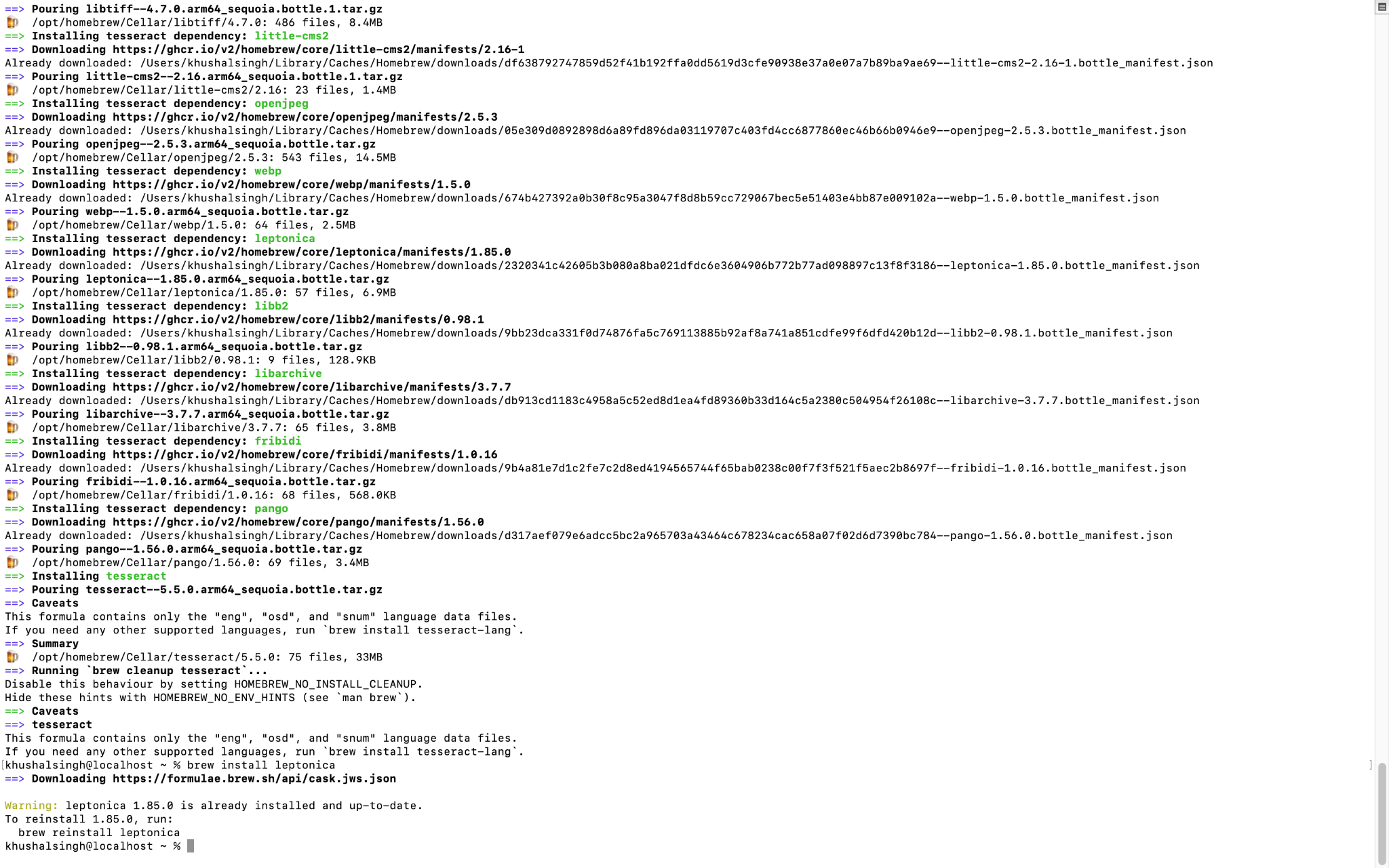
Nuget is present is present only for Tesseract not Terrasect

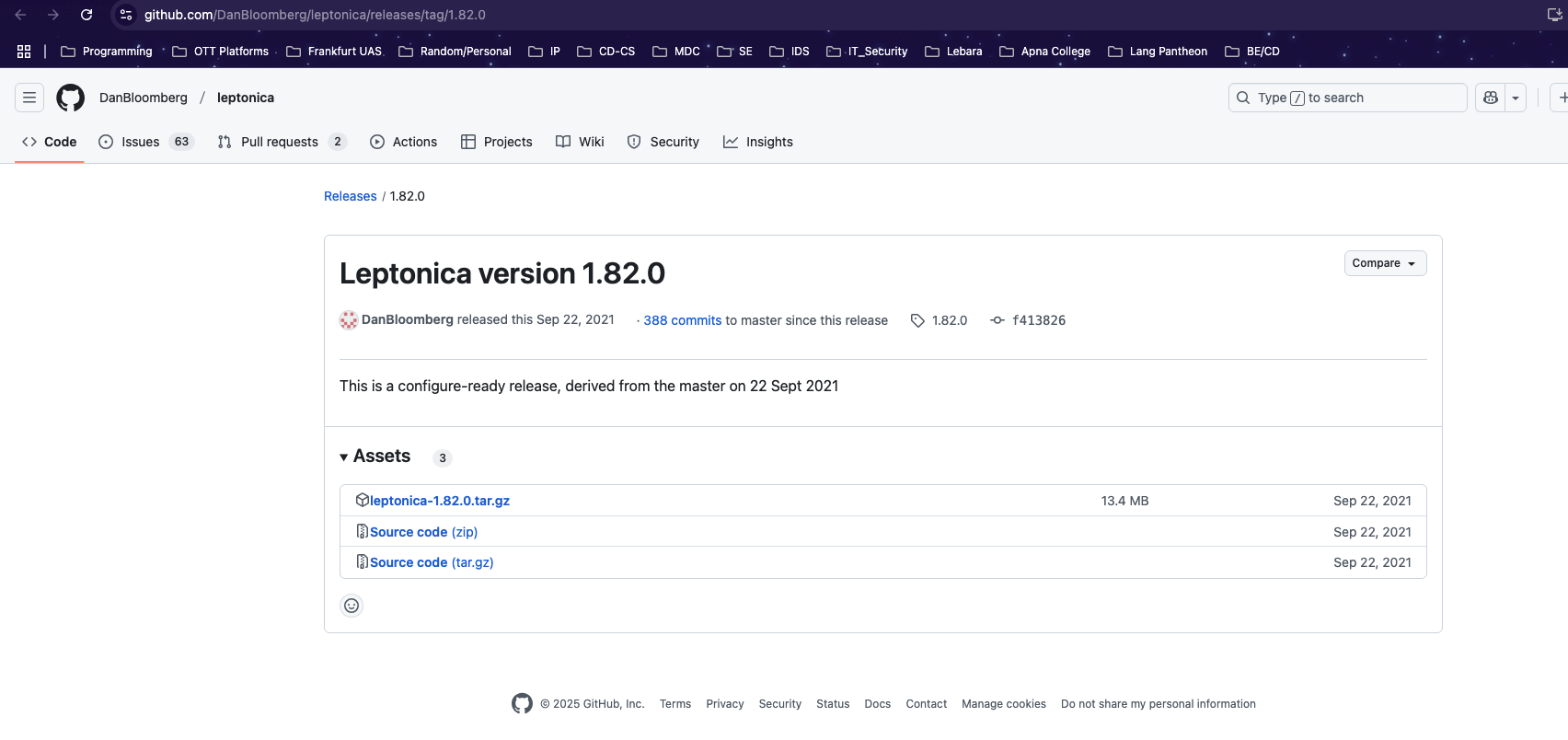


In the Document where project are described there is a potential mistake that Tesseract is replaced with Terrasect.

Homebrew Install



Install Tesseract and Leptonica   


Leptonica version 1.82.0 (tar.gz)  
****

Terminal   
khushalsingh@localhost ~ % cd Downloads

khushalsingh@localhost Downloads % cd leptonica-1.82.0

khushalsingh@localhost leptonica-1.82.0 % brew install pkg-config jpeg libpng libtiff webp

khushalsingh@localhost leptonica-1.82.0 % export DYLD\_LIBRARY\_PATH=/opt/homebrew/Cellar/leptonica/1.82.0/lib:$DYLD\_LIBRARY\_PATH

khushalsingh@localhost ~ % brew --prefix leptonica

x

khushalsingh@localhost leptonica-1.82.0 % brew install cmake

khushalsingh@localhost leptonica-1.82.0 % rm CMakeCache.txt

khushalsingh@localhost leptonica-1.82.0 % mkdir build

khushalsingh@localhost leptonica-1.82.0 % cd build

khushalsingh@localhost build % cmake ..

khushalsingh@localhost build % make

khushalsingh@localhost build % sudo make install

khushalsingh@localhost build % ls /usr/local/lib | grep leptonica

libleptonica.a

khushalsingh@localhost build % export DYLD\_LIBRARY\_PATH="/usr/local/lib:$DYLD\_LIBRARY\_PATH"

khushalsingh@localhost ConsoleApp1 % cd /Users/khushalsingh/Downloads/ocr-techtitans/src/OCR/OCRPreprocessor/Tesseract/ConsoleApp1/ConsoleApp1

khushalsingh@localhost ConsoleApp1 % dotnet build

khushalsingh@localhost ConsoleApp1 % dotnet add package System.Drawing.Common --version 6.0.0

khushalsingh@localhost ConsoleApp1 % tesseract /Users/khushalsingh/Downloads/test\_02.jpg /Users/khushalsingh/Downloads/ocr-techtitans/Output/output1.txt

khushalsingh@localhost ~ % brew info leptonica

If you need higher accuracy, you can download the **tessdata\_best** files. However, they are larger and slower.

If you need faster processing, use **tessdata\_fast** files, which are optimized for speed but with a slight trade-off in accuracy.

You can find both **tessdata\_best** and **tessdata\_fast** in the same repository.