

# Digit Recognition in Mobile Devices

Student: Jakub Kříž

Supervisor: doc. RNDr. Vlastislav Dohnal, Ph.D.

# Main points

- Thesis Objectives
- Goggles
- Android application – Environment
- OCR – Preprocessing, Feature Extraction, Classification
- OpenCV – Android, Manager, Examples, OpenCV Demo2, VLFeat
- Tesseract – tess-two, NDK, Cygwin, Character recognition, android-ocr
- Training Tesseract – Tesseract-OCR, Tesseract3, 3rdParty, fonts, images
- Image Preprocessing – manual, automatic, Leptonica

# Thesis Objectives

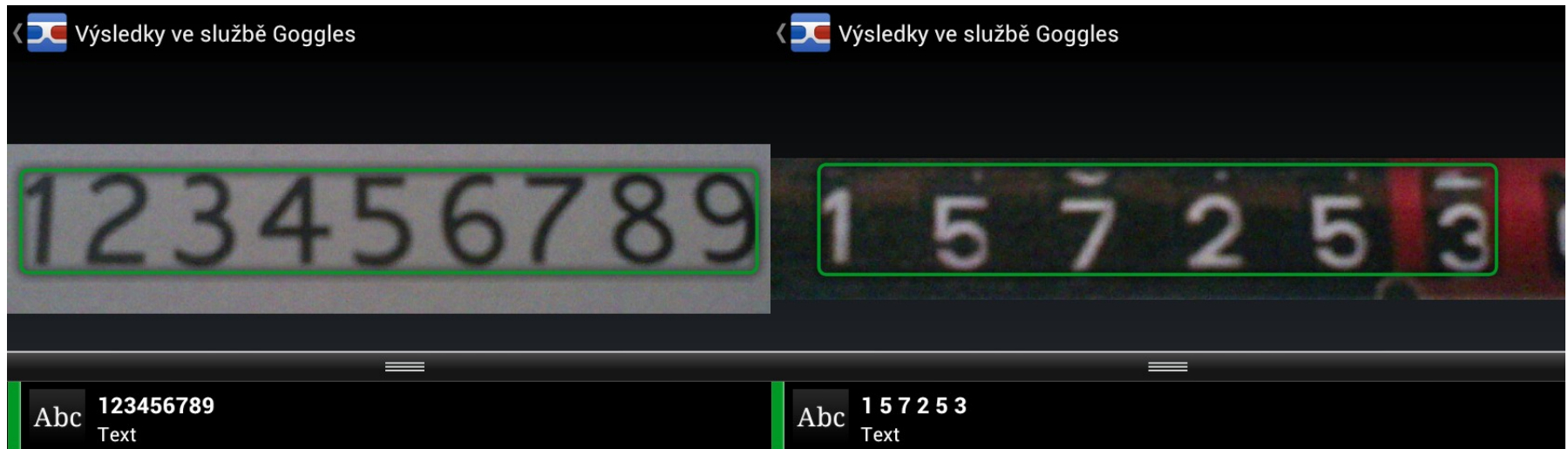
- Android Application – camera, box
- OCR – digit recognition
- High Success Rate
- Electric, Gas, Water Meters
- Extension – parts, Tomáš Lexmaul's diploma thesis



# Goggles

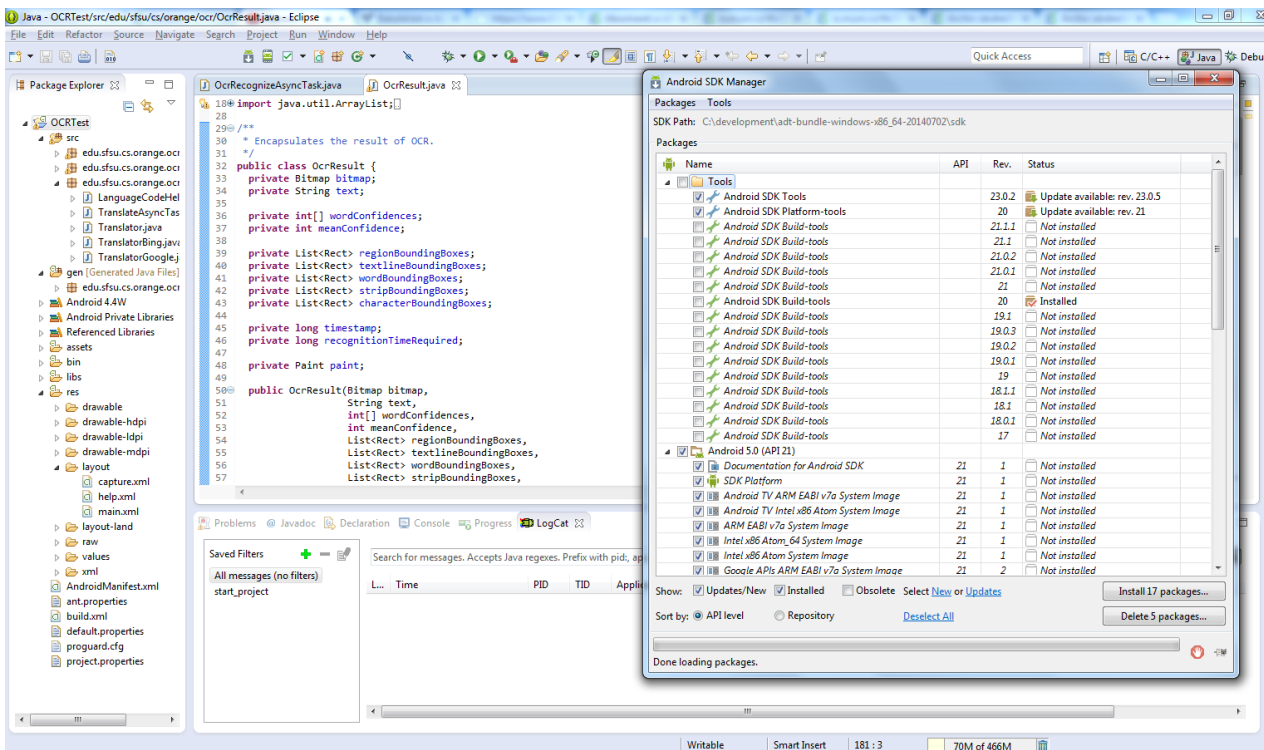
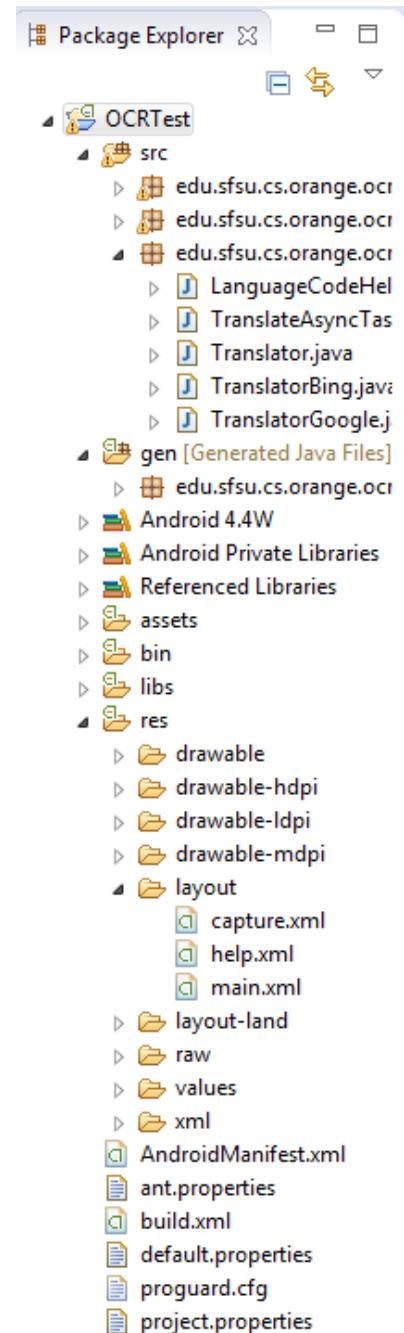


- Existing Solution – box, camera
- Internet
- Additional Functions (Search, Translation, Sudoku, QR, sightseeings)



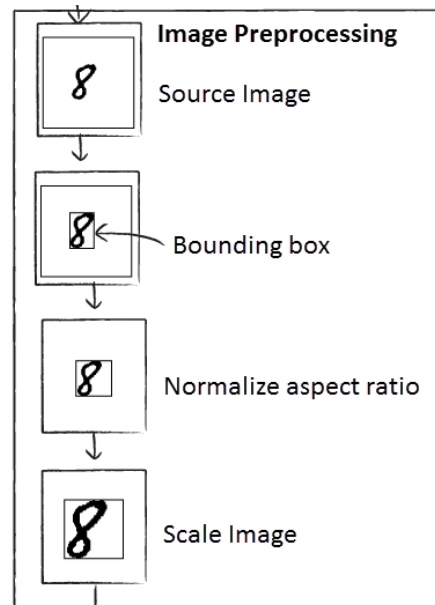
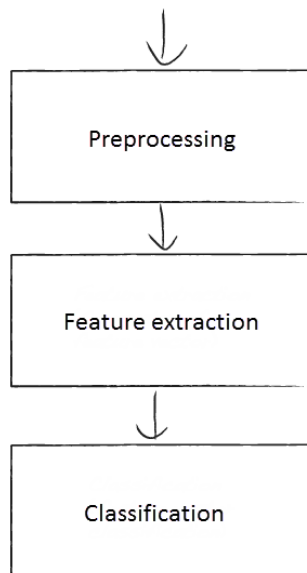
# Android Application

- [developer.android.com/training](http://developer.android.com/training)
- SDK (Eclipse ADT Bundle)
- SDK Manager
- Eclipse – Import... Existing Projects Into Workspace



# OCR

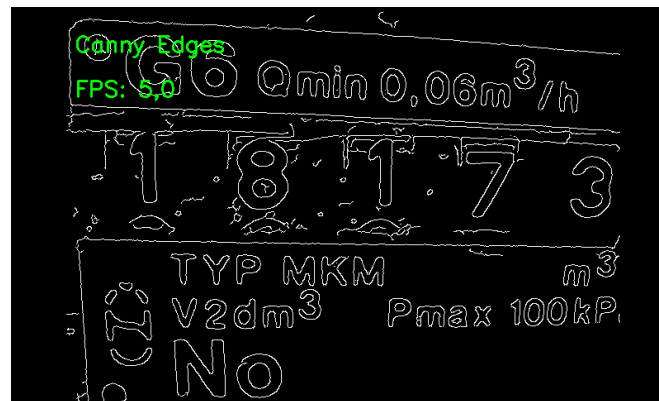
- [blog.damiles.com/2008/11/basic-ocr-in-opencv](http://blog.damiles.com/2008/11/basic-ocr-in-opencv)
- Preprocessing – input image, filtering, size normalizing, colour converting, bounding boxes, ...
- Feature extraction – image conversion, vector of features to classify
- Classification – feature vector, train system / classification method as knn
- my topic – 1 row, digits only, gaps, artefacts





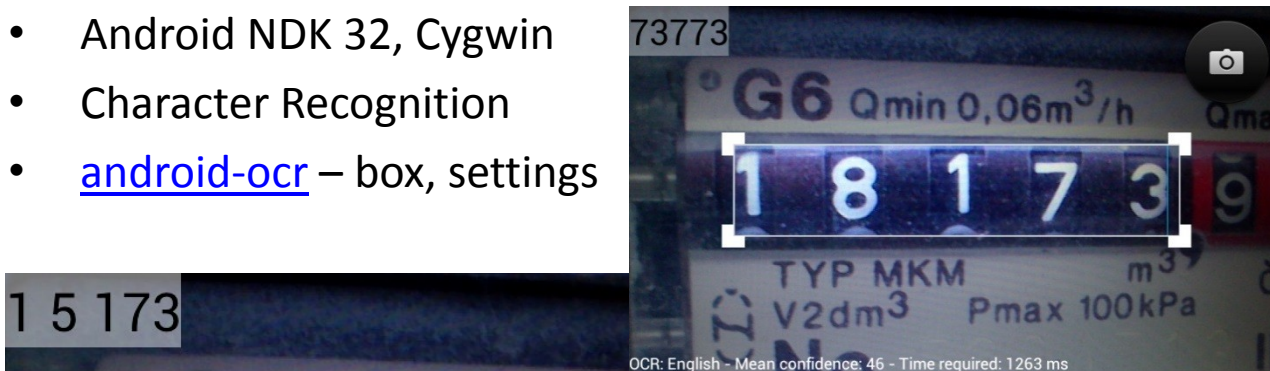
# OpenCV

- open source computer vision and machine learning software library
- 2500+ optimized algorithms
- detect and recognize faces, identify objects, classify human actions in videos, ...
- Android v2.4.9 – library as project, added as referenced project
- OpenCV Manager – needs to be installed along the app., does not require NDK
- examples - C++, Python, face recognition
- OpenCV Demo2 – Barry Thomas, source codes, Canny Edges, Track Features, ...
- VLFeat – above OpenCV, visual features (HOG), statistical methods (SVM)

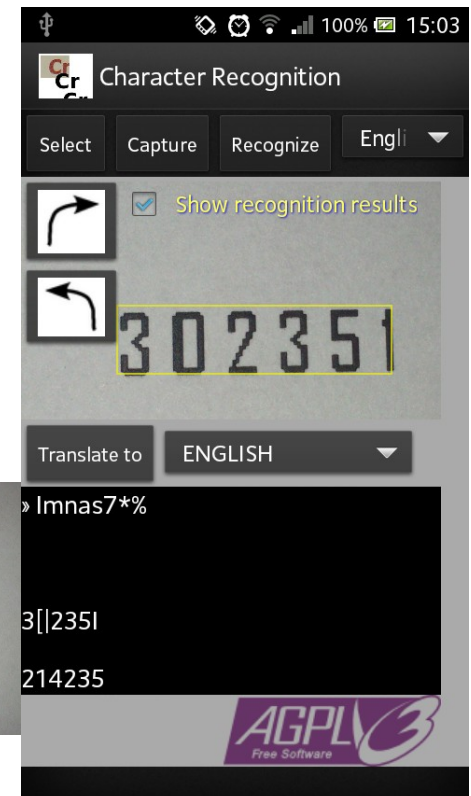


# Tesseract

- Open Source OCR engine, Apache 2.0 license
- Use - directly / API, languages, built-in GUI not provided
- `tesseract imagename outputbase [-l lang] [-psm pagesegmode] [configfile...]`
- `tess-two` (tesseract-android-tools + Leptonica )
- Android NDK 32, Cygwin
- Character Recognition
- [android-ocr](#) – box, settings

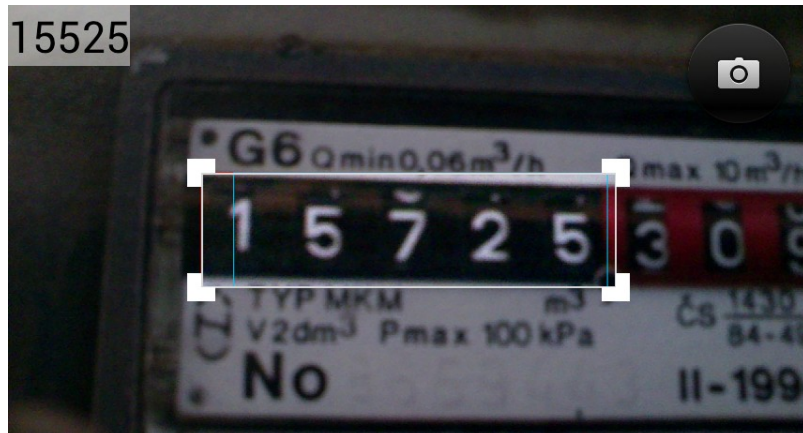


01711887  
302351  
214236



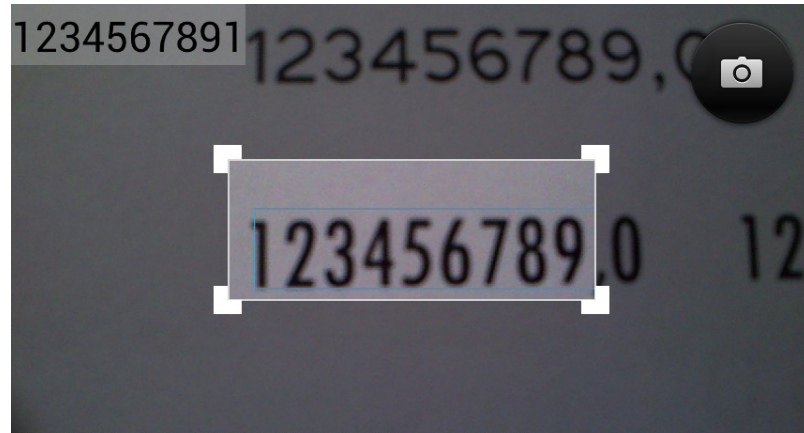


15525



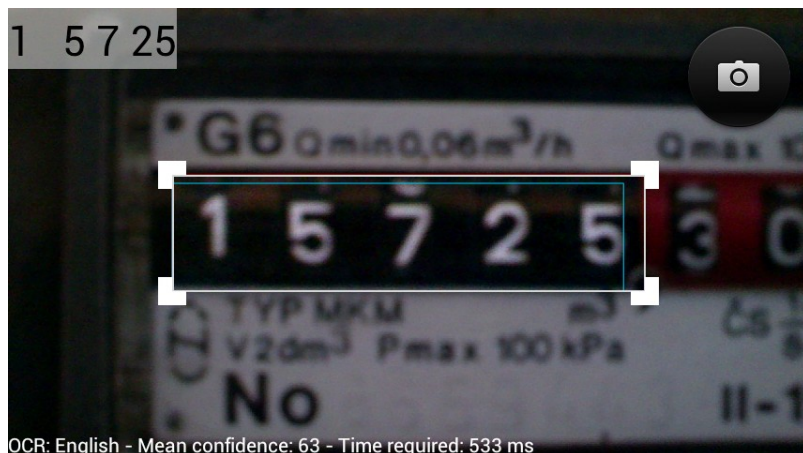
OCR: English - Mean confidence: 62 - Time required: 294 ms

1234567891



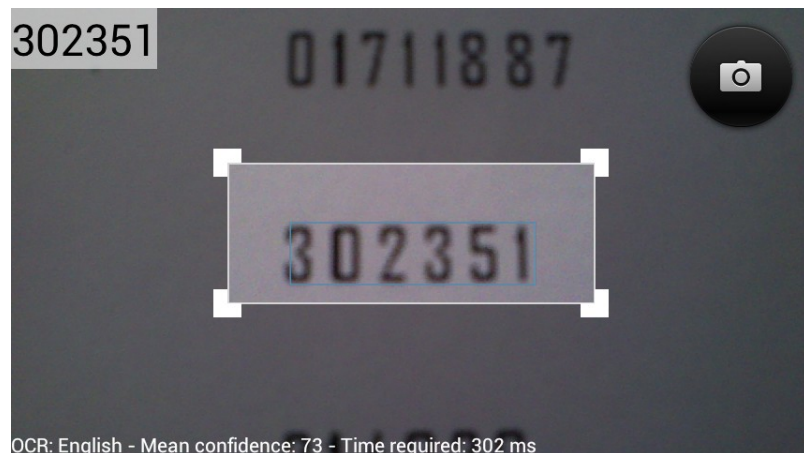
OCR: English - Mean confidence: 25 - Time required: 476 ms

1 5 7 25



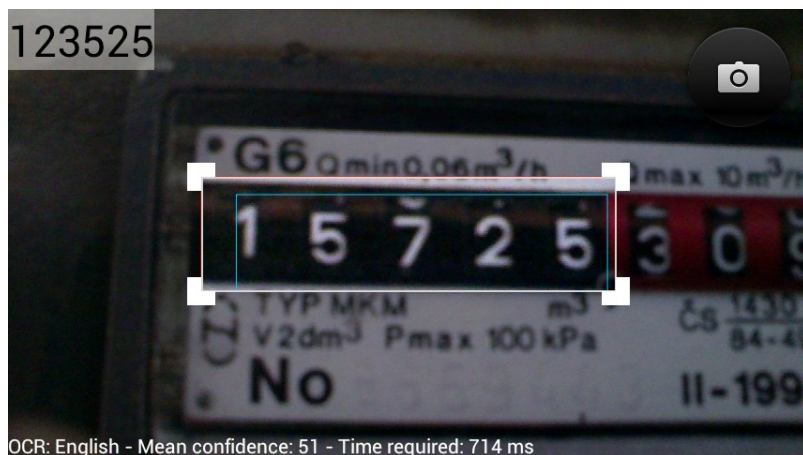
OCR: English - Mean confidence: 63 - Time required: 533 ms

302351



OCR: English - Mean confidence: 73 - Time required: 302 ms

123525



OCR: English - Mean confidence: 51 - Time required: 714 ms

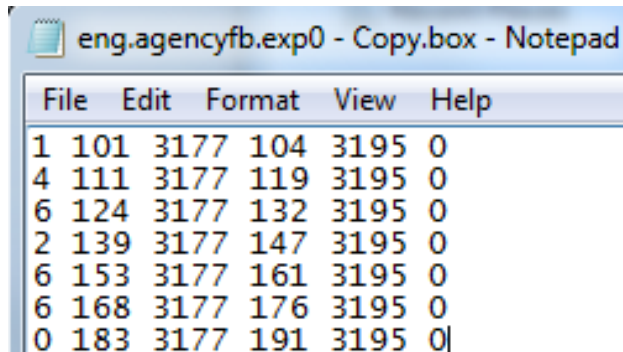
13173



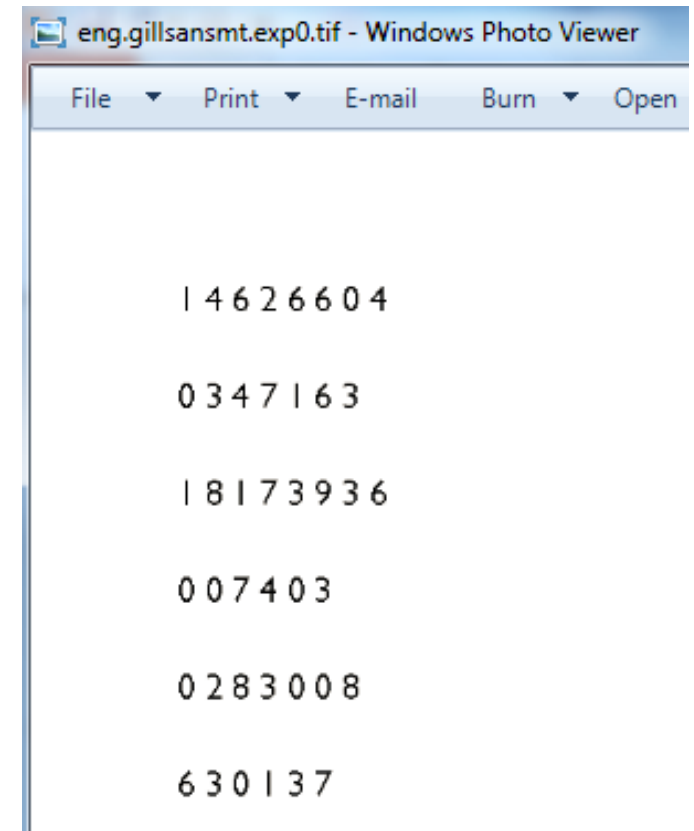
OCR: English - Mean confidence: 63 - Time required: 906 ms

# Training Tesseract

- Tesseract-OCR download
- Tesseract3 - box files ([0,0] bottom-left), train, font\_properties, combine
  - jTessBoxEditor, Serak
  - fonts vs. images (real data)
- Real data problem



```
eng.agencyfb.exp0 - Copy.box - Notepad
File Edit Format View Help
1 101 3177 104 3195 0
4 111 3177 119 3195 0
6 124 3177 132 3195 0
2 139 3177 147 3195 0
6 153 3177 161 3195 0
6 168 3177 176 3195 0
0 183 3177 191 3195 0
```



```
eng.gillsansmt.exp0.tif - Windows Photo Viewer
File Print E-mail Burn Open
1 4626604
0 347163
1 8173936
0 07403
0 283008
6 30137
```

<fontname> <italic> <bold> <fixed> <serif> <fraktur>  
timesitalic 1 0 0 1 0

Open Save Reload Merge Split Insert Delete

Character 4 X 119 Y 105 W 11 H 17

Box Coordinates Box Data Box View

```

1 105 3178 107 3195 0
4 119 3178 130 3195 0
6 138 3178 148 3195 0
2 157 3178 167 3195 0
6 176 3178 186 3195 0
6 194 3178 204 3195 0
0 213 3178 223 3195 0
4 231 3178 242 3195 0

```

1 4 6 2 6 6 0 4

Open Save Reload Merge Split Insert Delete

Character 8 X 354 Y 244 W 101 H 142

Box Coordinates Box Data Box View

	Char	X	Y	Width	Height
1	1	141	238	52	140
2	8	354	244	101	142
3	1	622	235	49	136
4	7	832	253	104	139
5	3	1074	251	98	136
6	9	1304	271	101	129
7	3	1532	268	96	134
8	6	1748	298	77	126



1. Add Train Image

2. Define Font\_Properties

3. Train Tesseract

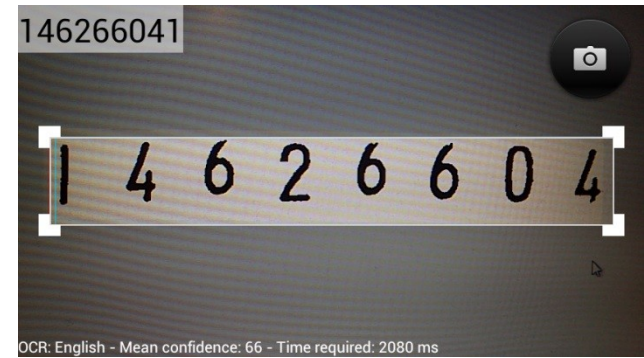
4. Combine TessData

Remove Selected

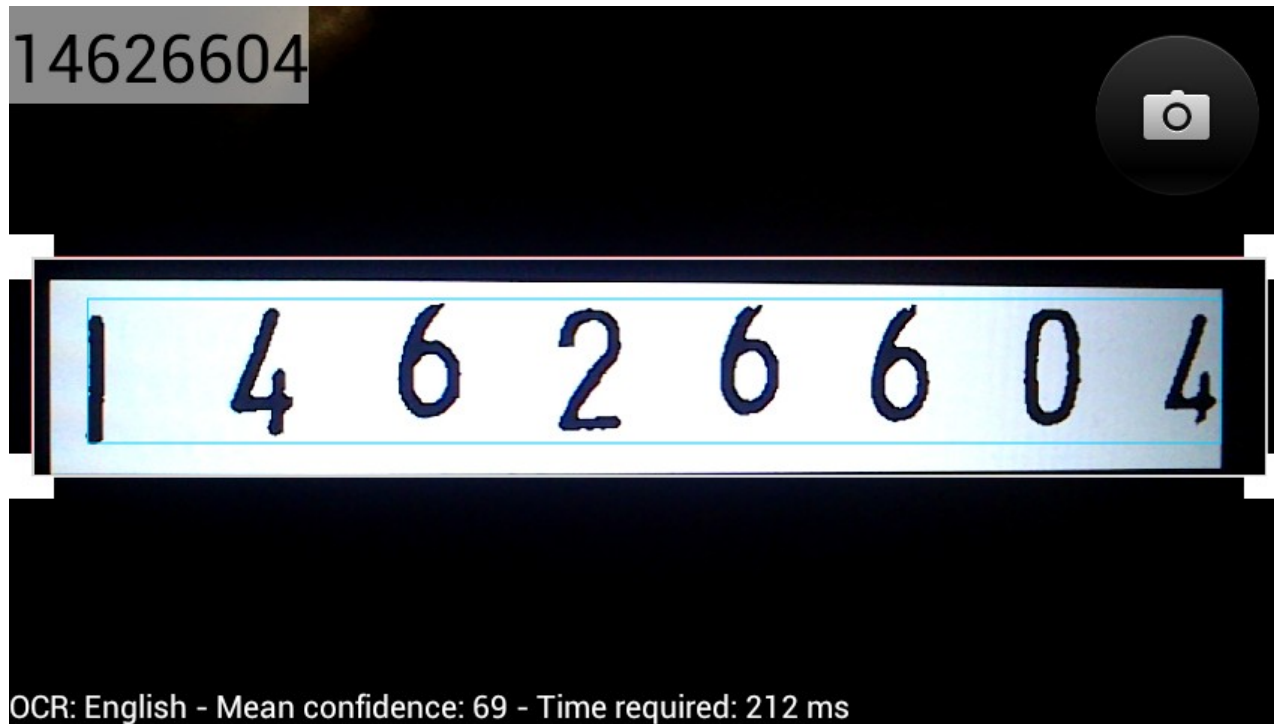
- eng.agencyfb.exp0.tif
- eng.bookmanoldstyle.exp0.tif
- eng.calibri.exp0.tif
- eng.candara.exp0.tif
- eng.couriernew.exp0.tif
- eng.euphemia.exp0.tif
- eng.forte.exp0.tif
- eng.gautami.exp0.tif
- eng.gillsansmt.exp0.tif
- eng.gillsansmtcondensed.exp0.tif
- eng.lilyupc.exp0.tif
- eng.lucidasans.exp0.tif
- eng.microsoftyibaiti.exp0.tif
- eng.miriamfixed.exp0.tif
- eng.moolboran.exp0.tif
- eng.raavi.exp0.tif
- eng.shruti.exp0.tif
- eng.simhei.exp0.tif
- eng.trebuchetms.exp0.tif
- eng.tunga.exp0.tif
- eng.twcenmtcondensed.exp0.tif

```
C:\development\training\real_data_training\TrainData - Copy>tesseract eng.myf
.exp0.tif eng.myfont.exp0 box.train
Tesseract Open Source OCR Engine v3.02 with Leptonica
row xheight=86, but median xheight = 69.5
row xheight=162.703, but median xheight = 69.5
row xheight=86, but median xheight = 69.5
row xheight=51, but median xheight = 69.5
row xheight=10, but median xheight = 69.5
row xheight=49, but median xheight = 69.5
row xheight=87, but median xheight = 69.5
FAIL!
APPLY_BOXES: boxfile line 0/1 <<141,468>>,<193,608>>: FAILURE! Couldn't find a
tching blob
FAIL!
APPLY_BOXES: boxfile line 1/8 <<354,460>>,<455,602>>: FAILURE! Couldn't find a
tching blob
FAIL!
APPLY_BOXES: boxfile line 2/1 <<622,475>>,<671,611>>: FAILURE! Couldn't find a
tching blob
FAIL!
APPLY_BOXES: boxfile line 3/7 <<832,454>>,<936,593>>: FAILURE! Couldn't find a
tching blob
FAIL!
APPLY_BOXES: boxfile line 4/3 <<1074,459>>,<1172,595>>: FAILURE! Couldn't find
matching blob
FAIL!
APPLY_BOXES: boxfile line 5/9 <<1304,446>>,<1405,575>>: FAILURE! Couldn't find
matching blob
FAIL!
APPLY_BOXES: boxfile line 6/3 <<1532,444>>,<1628,578>>: FAILURE! Couldn't find
matching blob
FAIL!
APPLY_BOXES: boxfile line 7/6 <<1748,422>>,<1825,548>>: FAILURE! Couldn't find
matching blob
APPLY_BOXES:
Boxes read from boxfile:      8
Boxes failed resegmentation: 8
APPLY_BOXES: Unlabelled word at :Bounding box=(0,760)->(2000,846)
APPLY_BOXES: Unlabelled word at :Bounding box=(0,480)->(942,760)
APPLY_BOXES: Unlabelled word at :Bounding box=(0,332)->(162,480)
APPLY_BOXES: Unlabelled word at :Bounding box=(692,332)->(942,480)
APPLY_BOXES: Unlabelled word at :Bounding box=(2,645)->(610,822)
APPLY_BOXES: Unlabelled word at :Bounding box=(641,645)->(1033,798)
APPLY_BOXES: Unlabelled word at :Bounding box=(573,644)->(733,654)
APPLY_BOXES: Unlabelled word at :Bounding box=(741,646)->(805,654)
APPLY_BOXES: Unlabelled word at :Bounding box=(815,645)->(989,654)
APPLY_BOXES: Unlabelled word at :Bounding box=(1040,649)->(1043,653)
APPLY_BOXES: Unlabelled word at :Bounding box=(1068,644)->(1123,653)
APPLY_BOXES: Unlabelled word at :Bounding box=(1137,643)->(1168,653)
APPLY_BOXES: Unlabelled word at :Bounding box=(1182,643)->(1238,651)
APPLY_BOXES: Unlabelled word at :Bounding box=(1279,686)->(1518,756)
APPLY_BOXES: Unlabelled word at :Bounding box=(1555,684)->(1854,810)
APPLY_BOXES: Unlabelled word at :Bounding box=(1917,630)->(2000,805)
APPLY_BOXES: Unlabelled word at :Bounding box=(1405,332)->(2000,760)
APPLY_BOXES: Unlabelled word at :Bounding box=(18,184)->(532,295)
APPLY_BOXES: Unlabelled word at :Bounding box=(676,214)->(904,277)
APPLY_BOXES: Unlabelled word at :Bounding box=(948,215)->(1252,280)
APPLY_BOXES: Unlabelled word at :Bounding box=(1585,223)->(1899,303)
APPLY_BOXES: Unlabelled word at :Bounding box=(1950,232)->(2000,294)
APPLY_BOXES: Unlabelled word at :Bounding box=(36,22)->(185,169)
APPLY_BOXES: Unlabelled word at :Bounding box=(230,62)->(454,189)
APPLY_BOXES: Unlabelled word at :Bounding box=(1420,0)->(2000,243)
Found 0 good blobs.
25 remaining unlabelled words deleted.
Generated training data for 0 words
```

# Image Preprocessing



- opened problem
- manual – Image Editor - grayscale, negative, manual threshold (remove artefacts)



- automatic – Leptonica - Grayscale, Inversion, Binarization, Thresholding

# Conclusion

- Thesis Objectives
- Goggles
- Android application
- OCR
- OpenCV
- Tesseract
- Training Tesseract
- Image Preprocessing

Thank you for your attention.