Picture Rotating and Labelling

Final Project Presentation

December 14th 2020
The Team

- Lennart Kitt (project owner, SEB)
- Arnel Pällo
- Karl-Kristjan Kõverik
- Sander Perens
- Oliver Ossip
The Problem

SEB gets different documents and real estate pictures that they need to

- Categorize
- Rotate upright
- Archive

Facade 90%

Document 60%
Acquired over 8000 photos from real estate portals
Preparing Training Data

- Removed borders and margins
- Labelled each one manually (gave up after 1000)
- Rotated all of them 90, 180 and 270 degrees (with a script)
Tested Different Models and Training Data

- Fast-AI with ResNet-34
- Build neural network ourselves with Keras
- Image-Net dataset
- One model with 12 classes vs several models with fewer classes
- Different image sizes
- Which augmentations improve performance?
- How big is our patience (can we tolerate training a model for 4 hours)?
Choosing the Best Approach

Confusion matrix

<table>
<thead>
<tr>
<th></th>
<th>document0</th>
<th>document180</th>
<th>document270</th>
<th>facade0</th>
<th>facade180</th>
<th>facade270</th>
<th>facade90</th>
<th>interior0</th>
<th>interior180</th>
<th>interior270</th>
<th>interior90</th>
</tr>
</thead>
<tbody>
<tr>
<td>document0</td>
<td>25</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>document180</td>
<td>25</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>document270</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade0</td>
<td>0</td>
<td>1</td>
<td>18</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade180</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade270</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior180</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior270</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Confusion matrix

<table>
<thead>
<tr>
<th></th>
<th>document0</th>
<th>document180</th>
<th>document270</th>
<th>facade0</th>
<th>facade180</th>
<th>facade270</th>
<th>facade90</th>
<th>interior0</th>
<th>interior180</th>
<th>interior270</th>
<th>interior90</th>
</tr>
</thead>
<tbody>
<tr>
<td>document0</td>
<td>10</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>document180</td>
<td>34</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>document270</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade180</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade270</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>facade90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior180</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior270</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>interior90</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Training data

Model

"It's a driver's license"

KITCHEN

247MemeGenerator.com
Results

- Best performance with Fast-AI using ResNet-34
- Used one model with 12 classes
- Model classified between documents, facades and interiors successfully
- Also classified 0 and 180 degree rotations successfully
- Confused at 90 vs 270 degrees
- Solution: if classified as rotated 90 or 270 degrees, rotate 90 degrees, then classify again. Finally rotate 180 degrees if needed
Architecture and workflow

- ipynb notebook for training and exporting the model (pickle)
- Command line script imports the model, then runs the classification and rotation
- Automated process in SEB systems.
  - Embed confidence in filenames of exported pictures
  - If model >= 90% confident, proceed
  - Else delegate for manual review
Next Steps

● Code committed to GitHub
  ○ https://github.com/perens/picture-labelling

● SEB engineers will take ownership
Lessons Learned

- Labelling is laborious. Maybe look for a career in unsupervised learning?
- Smaller images classify just as successfully, but train much faster…
- … except when it’s not the case (images of documents)
- Augmentation does not always help. It did not make our model more accurate
- FastAI worked wonderfully for us. Invest time into learning it!
- Some things don’t work and we have no idea why
- Some things do work but we have no idea why
Fast-AI

[classified as easy-peasy]
Thank you!

- SEB and Lennart Kitt for an interesting and very well defined problem
- Dima and TAs for a VERY well-organized and easy to follow course