Real-Time Automatic License Plate Recognition Through Deep Multi-Task Networks

Gabriel Resende Gonçalves
Matheus Alves Diniz
Rayson Laroca
David Menotti
William Robson Schwartz

Federal University of Minas Gerais
Federal University of Paraná
Outline

● Introduction
● ALPR Approach
  ○ Detection Net
  ○ Recognition Net
● Data Augmentation
● SSIG-ALPR Dataset
● Experimental Evaluation
● Conclusion
Introduction
Introduction

Automatic License Plate Recognition (ALPR) consists on perform on-track license plate recognition.

- **Key challenges**
  - Handle multiple vehicles
  - Execute on real-time
  - Predict correctly the majority of vehicles
Introduction

Usually, approaches divide license plate recognition into five subtasks and execute them in sequence:

- **a) image acquisition**
- **b) vehicle detection**
- **c) plate detection**
- **d) segmentation**
- **e) OCR**

**Drawback:** errors resulting of each task are propagated to the next step through the entire ALPR workflow.
Introduction

Contributions:

➔ A new public available ALPR dataset
➔ A new ALPR approach composed by two deep multi-task networks
➔ Three techniques to augment the training data

Hypothesis:

➔ We can overcome the error-rate propagation problem by performing ALPR with fewer tasks
➔ Some ALPR tasks such as character segmentation do not need to be explicit performed
ALPR Approach
ALPR Approach

Detection Network: detect on-road license plates directly on the frame

Recognition Network: recognize license plates with implicit segmentation

a) image acquisition

b) plate detection

c) plate recognition
ALPR Approach

Detection Net
ALPR Approach

Detection Net

Our loss penalizes regressions inside the license plate bounding box to ensure all characters will be completely visible.

a) underdetection

a) overdetection
ALPR Approach

Recognition Net
Data Augmentation
Data Augmentation

**Detection Net: Zoom**

- We need to train all anchors to ensure robustness for multiple scales

- **Solution:** Zoom-in and zoom-out the frames
Data Augmentation

Recognition Net: Character Permutation

- Every task of the proposed network has to learn the representation of each letter/digit
- Very hard due to the Brazilian license plate allocation policy

- **Solution**: Permute license plate characters
Data Augmentation

**Recognition Net:** Synthetic License Plates

- Since the permutations occur only between characters in the same plate, an undesired correlation between the characters in different positions was created

- **Solution:** Use synthetic license plates to train the fully-connected layers
SSIG-ALPR Dataset
SSIG-ALPR

Proposed Dataset

Our proposed dataset contains:

- 8,683 license plates images
- 815 different vehicles
- 3,368 images do not contain text annotation as they have very low resolution

2 cameras
Experiments
Experiments

Overview

- Detection Network Evaluation
- Recognition Network Evaluation
- Comparison with State-of-the-Art Approaches
  - **Baselines:** Silva and Jung [2], Gonçalves et al [3], Laroca et al [4], OpenALPR [5], Sighthound [6]
  - **Two datasets:** SSIG-SegPlate [3], UFPR-ALPR [4]
  - **Frame rate evaluation**

---

## Experiments

### Detection Network Evaluation

<table>
<thead>
<tr>
<th>Approach</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>no modification</td>
<td>76.73</td>
</tr>
<tr>
<td>loss only</td>
<td>+1.74</td>
</tr>
<tr>
<td>zoom only</td>
<td>+1.95</td>
</tr>
<tr>
<td>zoom + loss</td>
<td>+2.59</td>
</tr>
</tbody>
</table>
Experiments
Recognition Network Evaluation

<table>
<thead>
<tr>
<th>Approach</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>no augmentation</td>
<td>82.96</td>
</tr>
<tr>
<td>permutation only</td>
<td>+0.76</td>
</tr>
<tr>
<td>synthetic only</td>
<td>-33.43</td>
</tr>
<tr>
<td>permutation + synthetic</td>
<td>+2.64</td>
</tr>
</tbody>
</table>
Experiments
Comparison with State-of-the-Art Approaches

SSIG-SegPlate
- Static camera
- 2,000 images
- 101 vehicles

UFPR-ALPR
- Moving camera
- 4,500 images
- 150 vehicles
## Experiments

Comparison with State-of-the-Art Approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>SSIG-SegPlate (%)</th>
<th>UFPR-ALPR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva and Jung [2]</td>
<td>63.1</td>
<td>-</td>
</tr>
<tr>
<td>Gonçalves et al. [3]</td>
<td>81.8</td>
<td>-</td>
</tr>
<tr>
<td>Sighthound</td>
<td>73.1</td>
<td>-</td>
</tr>
<tr>
<td>OpenALPR</td>
<td>87.4</td>
<td>57.9</td>
</tr>
<tr>
<td>Laroca et al. [4]</td>
<td>85.4</td>
<td>72.2</td>
</tr>
<tr>
<td><strong>Proposed Approach</strong></td>
<td><strong>88.8</strong></td>
<td><strong>55.6</strong></td>
</tr>
</tbody>
</table>
## Experiments

### Comparison with State-of-the-Art Approaches: Frame Rate Evaluation

<table>
<thead>
<tr>
<th>Approach</th>
<th>Max # of license plates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva and Jung [2]</td>
<td>3</td>
</tr>
<tr>
<td>Gonçalves et al. [3]</td>
<td>1</td>
</tr>
<tr>
<td>Laroca et al. [4]</td>
<td>1</td>
</tr>
<tr>
<td><strong>Proposed Approach</strong></td>
<td><strong>6</strong></td>
</tr>
</tbody>
</table>
Final Remarks
Final Remarks

- **Data augmentation techniques are very helpful** to improve the network learning process.
- License plate detection **robustness is considerably diminished** when the images were acquired from a non-static cameras.
- In SSIG-SegPlate, our approach was able to **outperform all baselines** composed by multiple steps using static background.
- By creating two small networks, **we were able to run our approach with more 30 fps** even with 6 vehicles to recognize at the same time.

- As future works:
  - Jointly train both networks
  - Apply our approach with other license plates layouts
  - Adapt the network to work with motorcycles
Thank you

Contact us:

Smart Sense Laboratory
Universidade Federal de Minas Gerais
Belo Horizonte, Brazil

gabrielrg@dcc.ufmg.br
http://www.smartsense.dcc.ufmg.br