Tattoo Detection Based on CNN and Remarks on the NIST Database

1, 2Qingyong Xu, 1Soham Ghosh, 1Xingpeng Xu, 1Yi Huang, and 1Adams Wai Kin Kong (adamskong@ntu.edu.sg)

1School of Computer Science and Engineering, Nanyang Technological University, Singapore,
2Department of Computer, Nanchang University, China

Presented by

Soham Ghosh, (Undergraduate Student)
June 15, 2016
Current practices

- Face detectors
- Porn detectors
- Child porn detectors
Why are tattoos important?

• Tattoos are an important soft biometric trait
• Many people have tattoos: estimated 45 million Americans
• Tattoos have a lot of information for investigation.
Target Application

There is a need to build robust automated algorithms for tattoo detection.
Why do we need to detect tattoos?

• To search other criminals related to the case.

• In child sexual offense cases **120 TB images and videos data** should show **a lot of offenders**. If they have tattoos, they can be identified easily.

• Tattoo searching algorithms have been developed.
Why do we need to detect tattoos?
(Case 1: For further investigation, our target)

Seized computers → Tattoo detection → Tattoo matching
Why do we need to detect tattoos?
(Case 2: Tattoo database construction, mentioned in the NIST challenge)

Note: Non-tattoos are likely faces because currently law enforcement agencies collect face and tattoos in the process.
## Past work

<table>
<thead>
<tr>
<th></th>
<th>NIST Tatt-C</th>
<th>Heflin et al.</th>
<th>Wilber et al.</th>
<th>Our study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training Samples</strong></td>
<td>• Positive: 1349, Negative: 1000</td>
<td>• Total: 150</td>
<td>• Positive: 50, Negative: 800</td>
<td>• Positive: 5,740, Negative: 4,260</td>
</tr>
<tr>
<td><strong>Testing Samples</strong></td>
<td>• Positive: 1349, Negative: 1000</td>
<td>• Positive: 50, Negative: 500</td>
<td>Total: 100</td>
<td>• Positive: 5,740, Negative: 4,260</td>
</tr>
<tr>
<td><strong>Remarks</strong></td>
<td>• 5-fold cross-validation, Images from inner environments, Negative images are faces</td>
<td>Negative images were collected from dermatology forums and face databases</td>
<td>All positive images are butterfly.</td>
<td>• 5-fold cross-validation, No limit on positive and negative samples, Images collected from Flickr</td>
</tr>
<tr>
<td><strong>Techniques</strong></td>
<td>-</td>
<td>One class SVM</td>
<td>Exemplar Codes</td>
<td>CNN</td>
</tr>
</tbody>
</table>
NIST Tattoo Recognition Challenge

• To advance research and development into automated image-based tattoo recognition technology
  – identifying tattoos,
  – detecting region of interest,
  – matching visually similar or related tattoos using different types of non-tattoo imagery (e.g., scanned print and sketch),
  – matching similar tattoos from different subjects and
  – detecting tattoos from images

• The NIST challenge is open-book.
# Results of NIST Tattoo Detection Challenge

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Non-tattoo detection accuracy</th>
<th>Tattoo detection accuracy</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>French Alternative Energies and Atomic Energy Commission (CEA_1)</td>
<td><strong>98.8%</strong></td>
<td>93.2%</td>
<td>95.6%</td>
</tr>
<tr>
<td>Compass Technical Consulting (Compass)</td>
<td>38.6%</td>
<td>79.8%</td>
<td>62.2%</td>
</tr>
<tr>
<td>MITRE Corporation (MITRE 1)</td>
<td>75.0%</td>
<td>73.4%</td>
<td>74.1%</td>
</tr>
<tr>
<td>MITRE Corporation (MITRE 2)</td>
<td>94.8%</td>
<td>92.4%</td>
<td>93.4%</td>
</tr>
<tr>
<td>Morpho/MorphoTrak (MorphoTrak)</td>
<td>95.0%</td>
<td><strong>97.2%</strong></td>
<td><strong>96.3%</strong></td>
</tr>
</tbody>
</table>
Questions to be answered

1. Can CNN outperform the past winner of Tatt-C challenge?

2. How does the training database affect detection performance?

3. Is the NIST database suitable for our target application?
Convolutional Neural Network

Binary Classification
NIST Tattoo Recognition Challenge Dataset

Positive (1349)  
Negative (1000)
Results: NIST dataset

Accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compass</td>
<td>62.20%</td>
</tr>
<tr>
<td>MITRE 1</td>
<td>74.10%</td>
</tr>
<tr>
<td>MITRE 2</td>
<td>93.40%</td>
</tr>
<tr>
<td>CEA_1</td>
<td>95.60%</td>
</tr>
<tr>
<td>MorphoTrak</td>
<td>96.30%</td>
</tr>
<tr>
<td>CNN</td>
<td>98.80%</td>
</tr>
</tbody>
</table>
### Results: NIST dataset

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Non-tattoo detection accuracy</th>
<th>Tattoo detection accuracy</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEA_1</td>
<td>98.8%</td>
<td>93.2%</td>
<td>95.6%</td>
</tr>
<tr>
<td>Compass</td>
<td>38.6%</td>
<td>79.8%</td>
<td>62.2%</td>
</tr>
<tr>
<td>MITRE 1</td>
<td>75.0%</td>
<td>73.4%</td>
<td>74.1%</td>
</tr>
<tr>
<td>MITRE 2</td>
<td>94.8%</td>
<td>92.4%</td>
<td>93.4%</td>
</tr>
<tr>
<td>MorphoTrak</td>
<td>95.0%</td>
<td>97.2%</td>
<td>96.3%</td>
</tr>
<tr>
<td>CNN</td>
<td><strong>98.9%</strong></td>
<td><strong>98.7%</strong></td>
<td><strong>98.8%</strong></td>
</tr>
</tbody>
</table>

**Remark 1:** CNN is better than all the four participants in the NIST challenge.
Flickr Datasets

• Downloaded using Flickr API
• Four dataset sizes
  – Flickr2349
  – Flickr3.5K
  – Flickr5K
  – Flickr10K
• Same ratio of *positive:*negative (1.349:1)
• Datasets available at http://forensics.sce.ntu.edu.sg/
• These datasets are more similar to images in IT devices of suspects.
Flickr Datasets

**Positive** (keyword: *tattoo*)

**Negative** (keyword: *human, face*)
Results: Cross-dataset experiments

<table>
<thead>
<tr>
<th>TEST TRAIN</th>
<th>NIST</th>
<th>Flickr</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIST</td>
<td>98.81%</td>
<td>65.77%</td>
</tr>
<tr>
<td>Flickr</td>
<td>83.31%</td>
<td>78.29%</td>
</tr>
</tbody>
</table>

- Key observations
  - Accuracy drops significantly when the Flickr2349 dataset is used for testing.
  - Train NIST - Test Flickr performs the **worst**.
  - Train Flickr – Test NIST is better than Train Flickr – Test Flickr

**Remark 2:** NIST dataset is not suitable to train classifiers for our target application, detecting tattoos in IT devices of suspects.

**Remark 3:** Flickr dataset is much more challenging.
What causes the drop in accuracy?

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Non-tattoo detection accuracy</th>
<th>Tattoo detection accuracy</th>
<th>Accuracy difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Train NIST – Test NIST</td>
<td>98.70%</td>
<td>98.89%</td>
<td>-0.19%</td>
</tr>
<tr>
<td>2) Train NIST – Test Flickr</td>
<td>43.40%</td>
<td>82.36%</td>
<td>-38.96%</td>
</tr>
<tr>
<td>3) Train Flickr – Test NIST</td>
<td>74.40%</td>
<td>81.02%</td>
<td>-6.62%</td>
</tr>
<tr>
<td>4) Train Flickr – Test Flickr</td>
<td>70.10%</td>
<td>93.18%</td>
<td>-23.08%</td>
</tr>
</tbody>
</table>

Observations
- Detection accuracy for non-tattoos is much lower
- Discrepancy is largest for experiment 2 and 4.
What causes the drop in accuracy? (negative class)

- Negative (Flickr)
- Negative (NIST)
What causes the drop in accuracy? (positive class)

- Positive (Flickr)
- Positive (NIST)
Results: Flickr

Accuracy

Flickr(2349)  Flickr(3.5K)  Flickr(5K)  Flickr(10K)
Conclusions and suggestions

• Flickr images are more **challenging**
  – More diverse, hence closer to target application setting
• NIST database is **suitable for tattoo database construction**.
• NIST database is **not suitable** for the **target application**.
• **Large**, unconstrained dataset is needed
Suggestions

- For tattoo database construction, our **prisoner data collection system** may be a better solution. Tattoos and their accurate locations are collected at the same time.

“*A preliminary report on a full-body imaging system for effectively collecting and processing biometric traits of prisoners*, IEEE Symposium Series on Computational Intelligence, 2014.
Future Work

• Collecting a larger database
• Improving network architecture
Acknowledgments

• Thanks:
  – NIST for sharing the data.
  – Ngan, Mei Lee, NIST pointed out the error of the database link in the revised version.
  – Renaissance Engineering Programme, for financially supporting my conference trip.

Flickr database http://forensics.sce.ntu.edu.sg/.
THANK YOU