Promises and challenges of deep learning for remote sensing image exploitation

Christian Heipke
IPI - Institut für Photogrammetrie und GeoInformation
Leibniz Universität Hannover

President, ISPRS
Table of contents

- Introduction

- Deep learning
  - for images
  - for remote sensing

- Conclusions
Global ICT trends

Digitisation

Internet of Things, AI

Mobility

Autonomous driving

Industry 4.0

SDG goals

adapted from Volker Liebig, ESA 2016
ARTIFICIAL INTELLIGENCE

ZF establishes an AI research center

The automobile industry supplier wants to focus its activities in relation to artificial intelligence (AI) in Saarbrücken. One key area of...

more

© HAZ, 2.4.2019
Increased data volume
Deep learning
- the answer to automation needs?
Deep learning for image exploitation

... “learning” an input-output mapping from examples by machines *(supervised classification)* – not so new ...

- traditional classification:
  - Input Data → Cleverly-Designed Features → ML model

- deep learning:
  - Input Data → Deep Features → Learning model

  features and model learned together, mutually reinforcing each other
AI / ML / DL

- AI / DL ("data science") recently caught lots of attention in science and society alike – big data
  - „2nd best for everything“
  - hardware developments (in part. GPU)
  - abundant (training) data

- one of 10 breakthrough technologies in MIT Tech. Review 2013

- recently pushed by Internet companies
  - Google, Facebook, Microsoft, Baidu, …

- said to rival (even resemble?) human cognitive ability
  - victory of AlphaGo software against Go champion Lee Sedol (Oct.’17)
  - „super human“ performance in computer vision challenges
Classification (CV terminology), examples

- airplane
- automobile
- bird
- cat
- deer
- dog
- frog
- horse
- ship
- truck
Success story of DL/CNN in computer vision

ImageNet Large Scale Visual Recognition Comp. (INLSVRC):

- 1.2 Mio test images, 1000 classes
- One class label per image
- Ever increasing no. of layers

Classification: ImageNet Challenge top-5 error

- ILSVRC'15 ResNet: 3.57
- ILSVRC'14 GoogleNet: 6.7
- ILSVRC'14 VGG: 7.3
- ILSVRC'13: 11.7 (8 layers)
- ILSVRC'12: 16.4 (8 layers)
- ILSVRC'11: 25.8 (shallow)
- ILSVRC'10: 28.2 (shallow)

- AlexNet: 152 layers
Typical remote sensing data

... are CNNs any good for us also

???
Example I: Road extraction / refinement

- first CNN application using aerial images
- binary classification (road / no road)
- training data: OSM
  • incomplete
  • partly with geometric errors
- models for errors in training data, learning of the related model parameters

Example II: land cover / land use update

OA (LU) = 89.3 %
OA (LC) = 82.9 %

ISPRS Semantic Labelling Challenge

ISPRS Semantic Labeling Contest (2D): Results

Click here for a description of evaluation measures

Vaihingen: 2D Labelling challenge

All quality measures except for overall are F1 scores in [%] using the reference with eroded boundaries. Mouse over the column headings will display more information.

<table>
<thead>
<tr>
<th>Abbrev.</th>
<th>imp surf</th>
<th>building</th>
<th>low_veg</th>
<th>tree</th>
<th>car</th>
<th>Overall</th>
<th>Strategy</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVL_1</td>
<td>86.3</td>
<td>90.8</td>
<td>78.2</td>
<td>84.2</td>
<td>56.8</td>
<td>84.7</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>SVL_2</td>
<td>82.1</td>
<td>82.8</td>
<td>71.6</td>
<td>81.6</td>
<td>51.9</td>
<td>79.4</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>SVL_3</td>
<td>86.6</td>
<td>91.0</td>
<td>77.0</td>
<td>85.0</td>
<td>55.6</td>
<td>84.8</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>SVL_4</td>
<td>86.1</td>
<td>90.9</td>
<td>77.6</td>
<td>84.9</td>
<td>59.9</td>
<td>84.7</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>SVL_5</td>
<td>86.1</td>
<td>90.3</td>
<td>75.6</td>
<td>84.7</td>
<td>45.9</td>
<td>84.0</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>SVL_6</td>
<td>86.0</td>
<td>90.2</td>
<td>75.6</td>
<td>82.1</td>
<td>45.4</td>
<td>83.2</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>ADL_1</td>
<td>88.1</td>
<td>92.0</td>
<td>79.0</td>
<td>86.5</td>
<td>59.0</td>
<td>86.1</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>ADL_2</td>
<td>89.0</td>
<td>93.0</td>
<td>81.0</td>
<td>87.8</td>
<td>59.5</td>
<td>87.3</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>ADL_3</td>
<td>89.5</td>
<td>93.2</td>
<td>82.3</td>
<td>88.2</td>
<td>63.3</td>
<td>88.0</td>
<td>s</td>
<td>D</td>
</tr>
<tr>
<td>UT_May</td>
<td>84.3</td>
<td>88.7</td>
<td>74.5</td>
<td>82.0</td>
<td>9.0</td>
<td>81.8</td>
<td>s</td>
<td>D</td>
</tr>
</tbody>
</table>
Example III: Dense matching using CNNs


Conclusions- evaluation

- many other examples in the last few years (see e.g. review Zhu et al., 197 reference entries)
  - more flexible wrt. different object characteristics
  - strength: learning features, classifier not so important
  - key to good performance: network depth, no. of training data
  - has outperformed just about all classical algorithms by a large margin, provided enough training data is available
  - is about to do the same in geometric tasks

- no expert-free totally automatic processing chain, but integration with human capabilities

Conclusions - evaluation

• Open-source CNN implementations available
  – Tensorflow (Google):  https://www.tensorflow.org
  – CAFFE2 (Facebook)  https://caffe2.ai/
  – ...

• CNN has arrived in the commercial world
  – Descartes Lab,  https://www.descarteslabs.com/
  – Simularity,  simularity.com/solutions/satellite-anomaly-detection/
  – Orbital Insight,  https://orbitalinsight.com/
  – NVIDIA (computer games!),  https://www.nvidia.com/
  – ...
Conclusions - evaluation

• CNN is not magic - fundamentally a classifier
  – based on correlation between data sets
  – an enormous amount of unknowns to be estimated
    • lots of training data, long training times (days … weeks)
  – sensitive to
    • overfitting ("curse of dimensionality")
    • non-rep. training data (incorrect / biased / unbalanced / …)
    • user choices, incl. hyper-parameters
      – network architecture, (non-lin.) activation function
      – design of loss function (function to be optimised)
      – training parameters (weight initialisation, learning rate, drop out rate, "momentum", batch size, regularisation, …)
    – generalisation and prediction capabilities unclear
      • the system can’t learn what it never saw
Volvo admits its self-driving cars are confused by kangaroos

Swedish company’s animal detection system can identify and avoid deer, elk and caribou, but is yet to work against the marsupials’ movements

▲ Kangaroos are responsible for about 90% of collisions between vehicles and animals in Australia – although most are not serious. Photograph: Paul Kane/Getty Images
Conclusions – to do’s

• integrate prior knowledge, e.g. physical models
  – don’t try to learn what we already know (e.g. laws of nature)

• CNNs for point clouds and other irregularly rep. data

• tackle training data shortage
  – weakly, semi-, unsupervised learning, reinforcement learning
  – data augmentation and simulation
  – deal with errors in training data, e.g. when using outdated DB
  – combination with crowd sourcing and social media data
  – investigate limits of pre-training + fine tuning

• pay attention to geometric accuracy of object delineation

• sequential data, time series processing - RNNs
  – streaming (big) data, online CNN learning
Special thanks to

Prof. Franz Rottensteiner
Chair, IPI research group
Photogrammetric image analysis

Chen Lin, Max Coenen, Philo Humberg, Uyen Nguyen, Dennis Wittich, Chun Yang, Junhua Kang
ISPRS Geospatial Week 2019

Enschede (NL), June 10 – 14, 2019

Deadline for full papers: Jan. 17, 2019
ISPRS Congress 2020
Nice, France
Imaging today, foreseeing tomorrow
www.isprs2020-nice.com
June 14 – 20, 2020