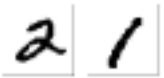
Hands-On Deep Learning for Computer Vision Projects

Dr. Leal-Taixe
Caner Hazirbas, Vladimir Golkov
Technische Universität München

What's a project?

- **there is a problem** handwritten digit recognition
- **the problem is relevant** automatically scan checks
- **the goal is clear**  → 21
- **there is a ML approach** deep neural net
- **there is a group** (= 2 people)

What do we want in the end?

- final presentation
  - what is the problem
  - why is the problem relevant
  - what is the goal
  - what is the approach
  - which experiments did you run (params, results)
  - what went well / what didn't
  - what's the state of the art
  - what can be done in the future

How to register

- pick 3 projects and rank them by preference
- send an email to dlpractice@vision.in.tum.de
- you can either form a group of 2 and send the group information. Or just send your preference and then we will form students into groups.
- please register before Oct 31, 2016, 12 a.m.

Projects

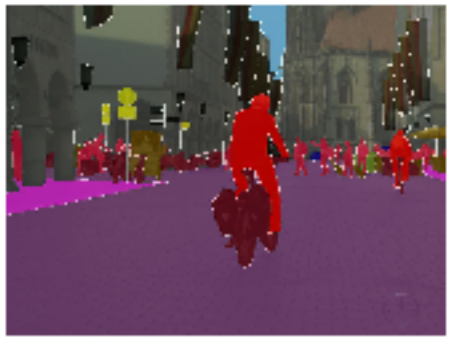
Instance-based detection/segmentation

- Implementation in TensorFlow
- **Goal:** detect and follow objects through a video
  No LSTM! Taking inspiration from Re-ID.

Projects

Instance-based detection/segmentation

- Implementation in TensorFlow
- **Goal:** detect and follow objects through a video
  No LSTM! Taking inspiration from Re-ID.
- There is room for **semantics**!

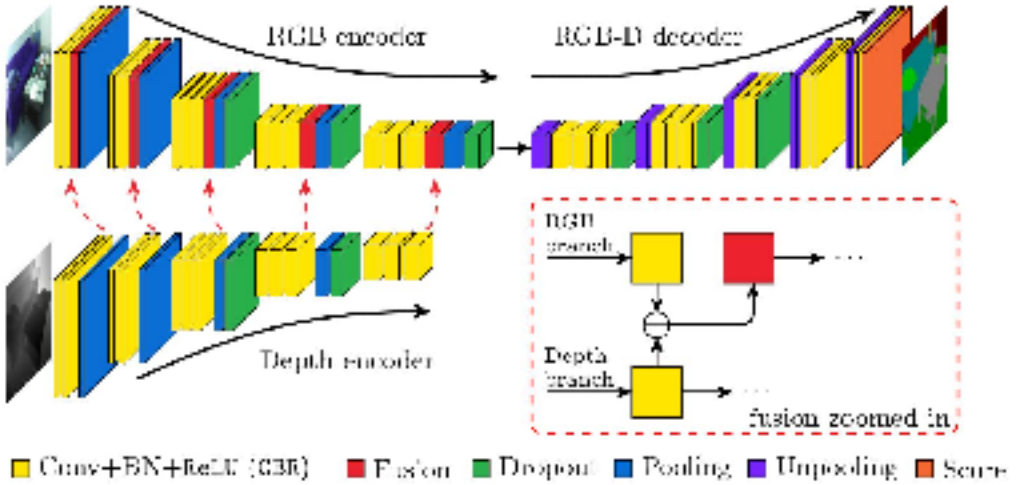


Healthy competition

Projects

Generalizing ***FuseNet*** to other applications and networks

- Implementation in TensorFlow
- **Group 1:** Scene classification & segmetation
- **Group 2:** Different network architectures

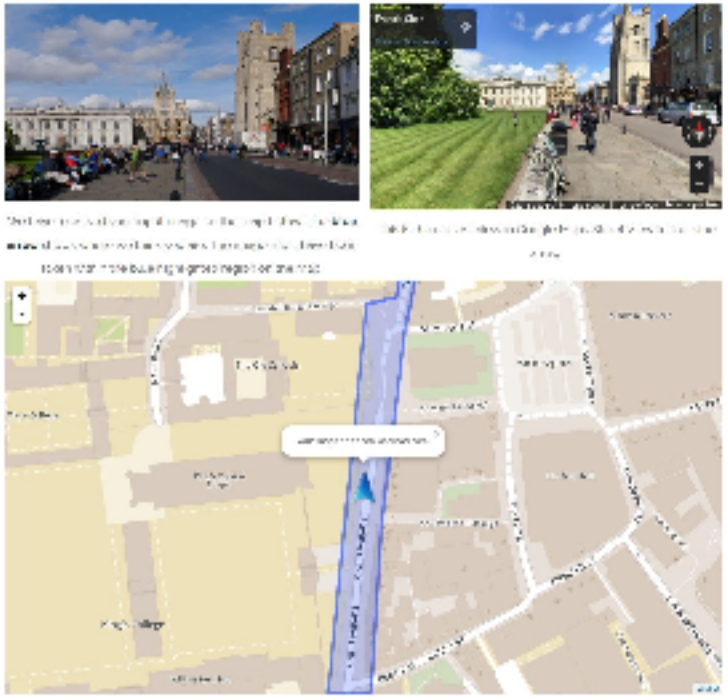


C. Hazirbas, L. Ma, C. Domokos, D. Cremers, *In Asian Conference on Computer Vision (ACCV)*, 2016

Projects

Video-based virtual localization (extension of PoseNet)

- Implementation in TensorFlow
- **Group 1:** using CNN
- **Group 2:** using LSTM



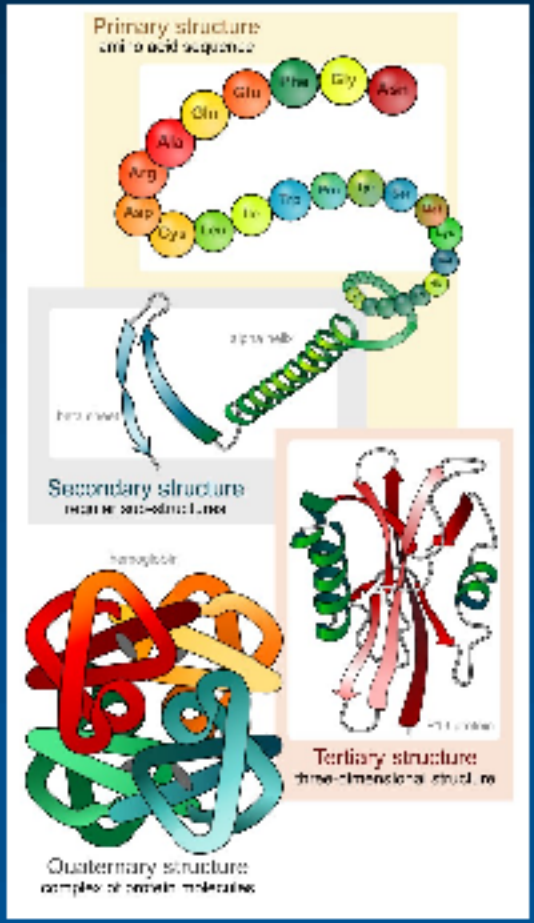
http://mi.eng.cam.ac.uk/projects/relocalisation/
Alex Kendall, Matthew Grimes and Roberto Cipolla "**PoseNet**: A Convolutional Network for Real-Time 6-DOF Camera Relocalization."

Introduction: Proteins

- "Building blocks of life"
- Structure and function important for life sciences

- Protein structure:
  - Primary structure: 1D sequence of amino acids
  - Secondary structure: local folding
  - Tertiary structure: global 3D shape
- Representation can also be 2D

- Working with these 1D, 2D, 3D representations is mutually beneficial for other fields (sequences, images, shapes…)



Projects

Foundations of Deep Learning

Several sub-projects such as:

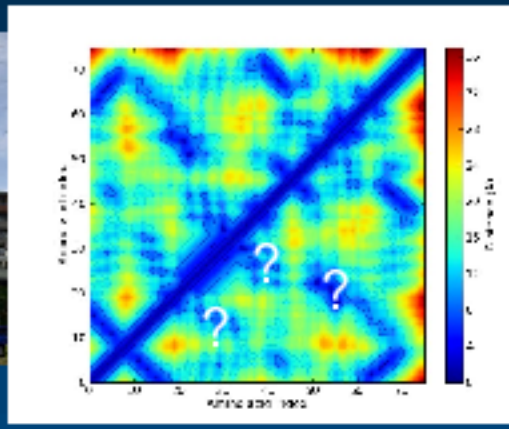
- Denoising of protein scoring matrices
  - 1D convolutional network
- Positive/unlabeled (PU) learning for virtual screening
  - Multilayer perceptron
  - Interesting cost function and setting
  - Dealing with data imbalance
- Great to get started, and to get intuition of data distributions and problem-tailored solutions

Improving Results of Different Networks

- The art of pushing the quality
- State-of-the-art methods
  - Deep architectures
  - All-convolutional networks
  - Residual learning
  - Densely connected convolutional networks
  - Advanced cost functions
  - And more
- Applications:
  - Super-resolution of cryo–electron microscopy (3D ConvNets)
  - Virtual screening (multi-input)
  - …

Recognition of Protein Structural Motifs in 2D Data

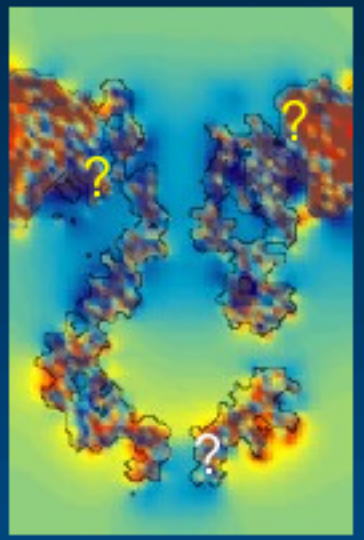
- 2D convolutional networks
- Classification, localization, weakly-supervised learning, PU learning



- Experience with array programming in NumPy (or Matlab or similar) is beneficial
- References:
  - http://docs.scipy.org/doc/numpy/reference/arrays.indexing.html
  - optional: Oquab et al. 2015: “Is object localization for free? – Weakly-supervised learning with convolutional neural networks”

Recognition of Protein Structural Motifs in 3D Data

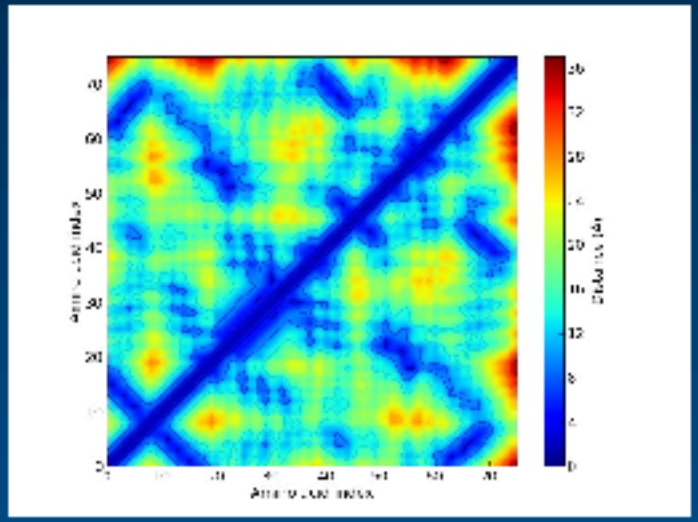
- 3D convolutional networks
- Classification, localization, weakly-supervised learning, PU learning
- Big data!



- Experience with array programming in NumPy (or Matlab or similar) is beneficial
- References:
  - http://docs.scipy.org/doc/numpy/reference/arrays.indexing.html
  - optional: Oquab et al. 2015: “Is object localization for free? – Weakly-supervised learning with convolutional neural networks”

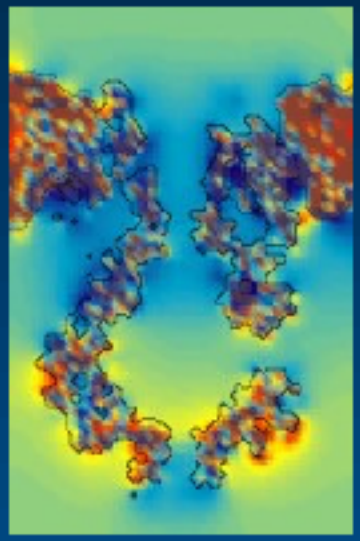
Protein Function Prediction from 2D Data

- 2D convolutional networks
- Interesting output structure (class hierarchy)
- Outlook: implementation of interesting biological features



- Interest in biology is beneficial
- Experience with array programming in NumPy (or Matlab or similar) is beneficial
- References:
  - http://docs.scipy.org/doc/numpy/reference/arrays.indexing.html

Protein Function Prediction from 3D Data



- 3D convolutional networks
- Interesting output structure (class hierarchy)
- Challenging task, advanced deep learning

- Requirements:
  - Practical experience with deep learning
  - Good skills in NumPy (or Matlab or similar)
- References:
  - http://docs.scipy.org/doc/numpy/reference/arrays.indexing.html

The Future of Medical Imaging

- “Image in, diagnosis out”
- Challenge: overfitting
- 3D convolutional networks
- Data: primarily Diffusion MRI