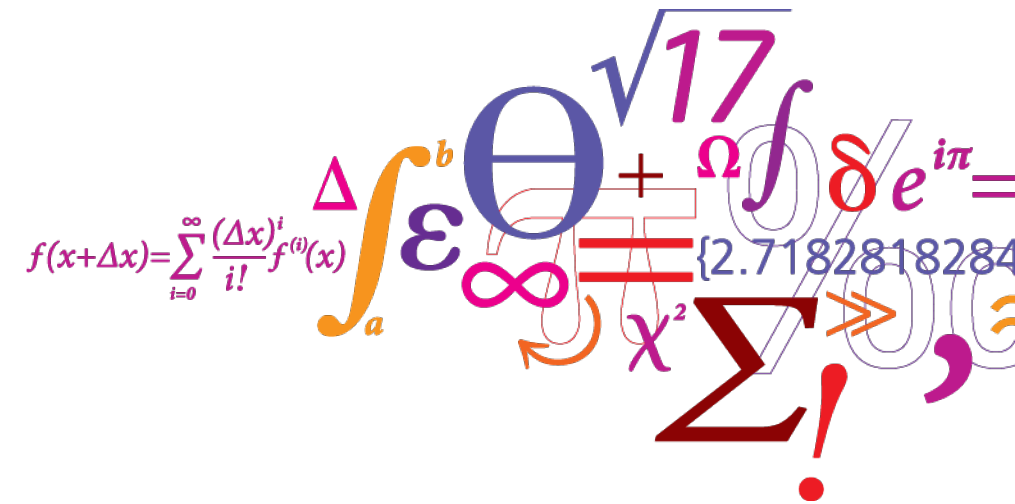


# 02460

Image Segmentation

Tommy Sonne Alstrøm

Jan Larsen



DTU Compute

Department of Applied Mathematics and Computer Science

# Perspective



IDUN is a center of excellence funded by the Danish National Research Foundation and the Villum Foundation. The center is divided into two parts: IDUN Drug and IDUN Sensor, focusing on drug delivery and nanomechanical sensors, respectively.



<http://www.idun.dtu.dk/>

# Project setup

Use <https://lab.compute.dtu.dk/> as repository

Weekly logbook – (can use wiki)

git clone [git@lab.compute.dtu.dk:02460F18-Image/shared.git](https://lab.compute.dtu.dk:02460F18-Image/shared.git)

git clone [git@lab.compute.dtu.dk:02460F18-Image/group1.git](https://lab.compute.dtu.dk:02460F18-Image/group1.git)

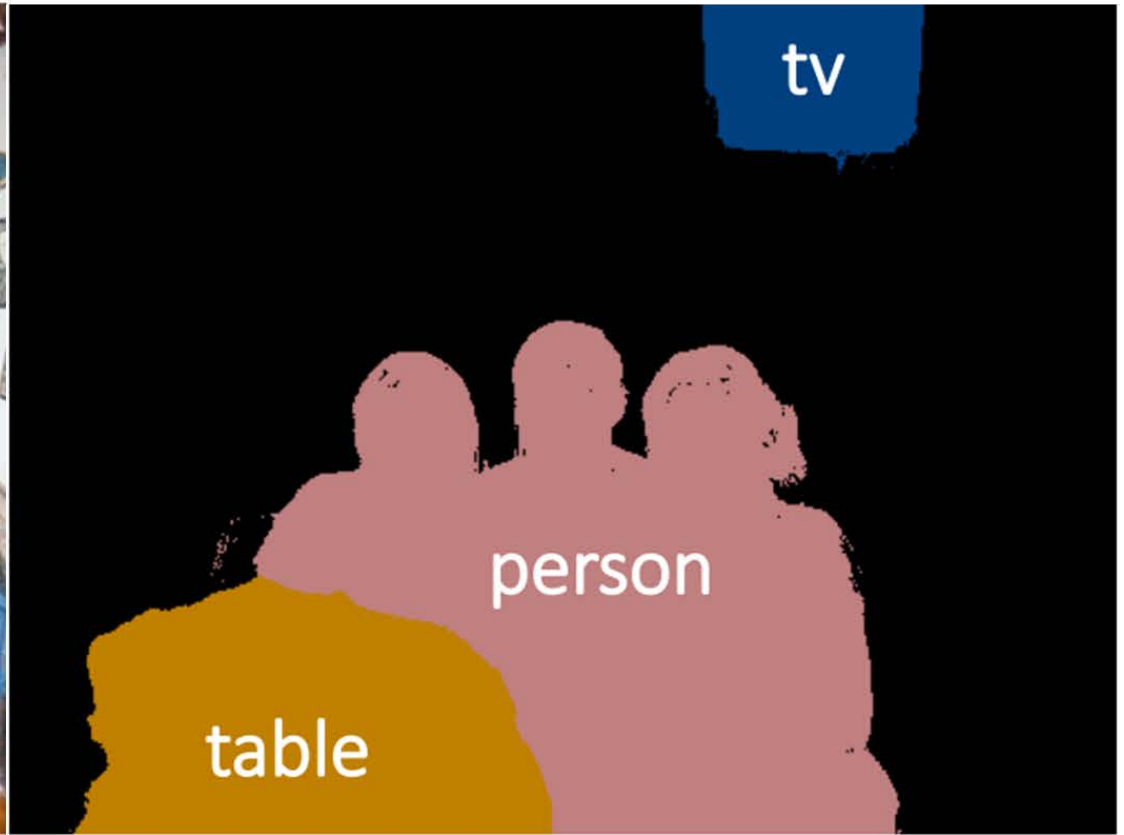
Future meetings 321/119 Thursdays 14.00 to 15.00.

# Learning objectives

- 1) Define own learning objectives for the project
- 2) Collect scientific knowledge and data related to the project topic based on a specific project proposal
- 3) Carry out a well-founded delimitation of the project and formulate specific hypotheses and aims
- 4) Organize and coordinate the work in the project group
- 5) Plan and carry out the course of the project in collaboration with the project supervisor
- 6) Design a machine learning based system starting from analysis of the problem and the project aims, and further select relevant algorithms and methods
- 7) Assess and summarize the project results in relation to aims, methods and available data
- 8) Carry out the project and interpret results by use of Matlab, Python or other programming language
- 9) Structure and write a final short technical article including problem formulation, description of methods, experiments, evaluation and conclusion
- 10) Presentation of methods and results at meetings with project supervisor and fellow students
- 11) Organize and present project results at the final poster presentation

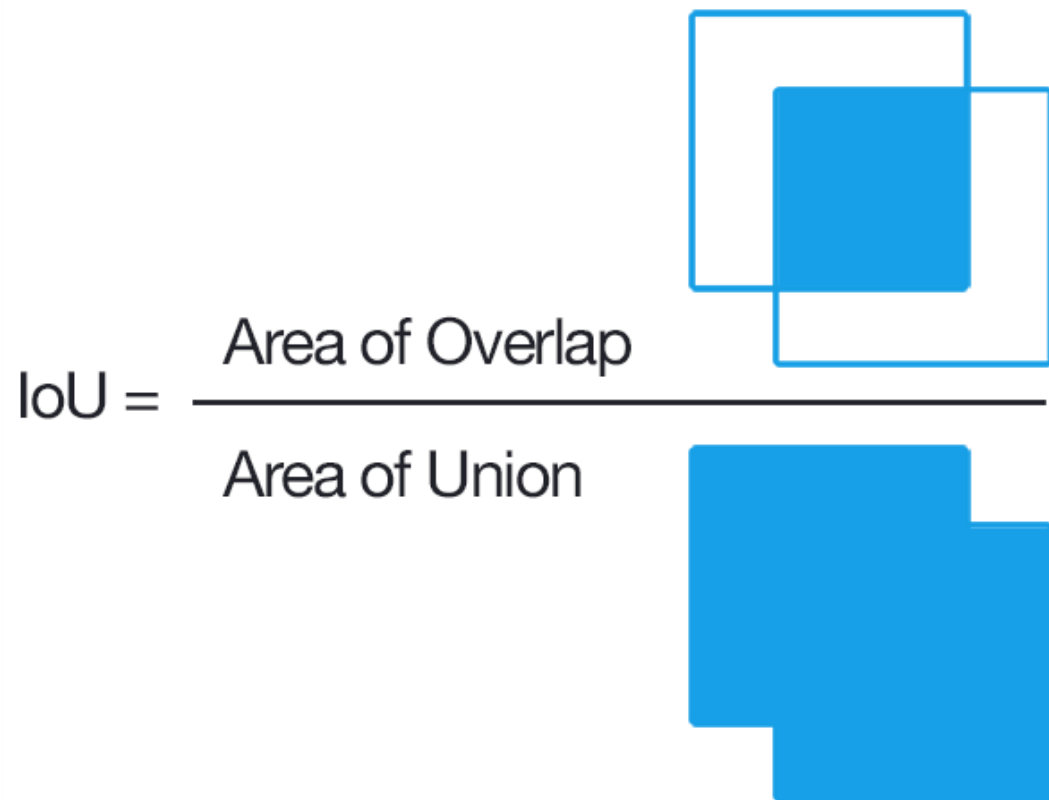
# Image segmentation

<https://github.com/kjw0612/awesome-deep-vision#semantic-segmentation>




# Evaluation – Intersection over Union (IoU) for object detection

<http://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>



# MIT indoor dataset

<http://web.mit.edu/torralba/www/indoor.html>



The image shows a grid of 100 small indoor scene images, categorized into five groups: Store, Home, Public spaces, Leisure, and Working place. Each group contains a 4x5 grid of representative images.

## Indoor Scene Recognition

Indoor scene recognition is a challenging open problem in high level vision. Most scene recognition models that work well for outdoor scenes perform poorly in the indoor domain. The main difficulty is that while some indoor scenes (e.g. corridors) can be well characterized by global spatial properties, others (e.g., bookstores) are better characterized by the objects they contain. More generally, to address the indoor scenes recognition problem we need a model that can exploit local and global discriminative information.

### Database

The database contains 67 indoor categories, and a total of 15620 images. The number of images varies across categories, but there are at least 100 images per category. All images are in jpg format. The images provided here are for research purposes only.

[Download \(tar file, 2.4 Gbytes\)](#)

### Evaluation

For the results in the paper we use a subset of the dataset that has the same number of training and testing samples per class. The partition that we use is:

- [TrainImages.txt](#): contains the file names of each training image. Total 67\*80 images
- [TestImages.txt](#): contains the file names of each test image. Total 67\*20 images

### Annotations

A subset of the images are segmented and annotated with the objects that they contain. The annotations are in LabelMe format.

[Download annotations](#)

### Paper

A. Quattoni, and A. Torralba. [Recognizing Indoor Scenes](#). IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

### Acknowledgments

Thanks to Aude Oliva for helping to create the database of indoor scenes.  
Funding for this research was provided by NSF Career award (IIS 0747120)




# Describable textures dataset

<https://www.robots.ox.ac.uk/~vgg/data/dtd/>

Downloads

DTD
Overview
Citation
Download
Browse
Releases
Evaluation

## Describable Textures Dataset (DTD)



Download dataset
Download code
Evaluation
Citation

### Overview

Our ability of vividly describing the content of images is a clear demonstration of the power of human visual system. Not only we can recognise objects in images (e.g. a cat, a person, or a car), but we can also describe

Filename	Description	Size
<a href="#">README.txt</a>	README file describing: <ul style="list-style-type: none"> <li>Dataset structure.</li> <li>Ground truth annotations: key attributes and joint attributes.</li> </ul>	185K
<a href="#">dtd-r1.0.1.tar.gz</a>	The package contains: <ul style="list-style-type: none"> <li>Dataset images, train, validation and test.</li> <li>Ground truth annotations and splits used for evaluation.</li> <li>imdb.mat file, containing a struct holding file names and ground truth labels.</li> </ul>	625M
<a href="#">dtd-r1.0.1-labels.tar.gz</a>	Annotations and splits <ul style="list-style-type: none"> <li>Ground truth annotations: key attributes, joint attributes.</li> <li>Splits of the data into train, val and test, as used in our experiments.</li> </ul>	1.4M
<a href="#">dtd-r1-decaf_feats.tar.gz</a>	Compressed decaf_feats.mat, containing a 5640x4096 matrix, represented DeCAF features for the images from DTD. Each row represents the 4096 dimensional feature vector for one image, assuming images are sorted by name.	82M

### Acknowledgements

This research is based on work done at the 2012 CLSP Summer Workshop, and was partially supported by NSF Grant #1005411, ODNI via the JHU-HLTCE and Google Research. Mircea Cimpoi was supported by the ERC grant VisRec no. 228180 and Iasonas Kokkinos by ANR-10-JCJC-0205.

The development of the describable textures dataset started in June and July 2012 at the Johns Hopkins Centre for Language and Speech Processing (CLSP) Summer Workshop. The authors are most grateful to Prof. Sanjeev Khudanpur and Prof. Greg Hager.



# Pascal VOC datasets

- <http://host.robots.ox.ac.uk/pascal/VOC/>

## The PASCAL Visual Object Classes Homepage



---

### Development Kit

The development kit consists of the training/validation data, MATLAB code for reading the annotation data, support files, and example implementations for each competition.

The development kit is now available:

- Download the [training/validation data](#) (2GB tar file)
- Download the [development kit code and documentation](#) (500KB tar file)
- Download the [PDF documentation](#) (500KB PDF)
- Browse the [HTML documentation](#)
- View the [guidelines](#) used for annotating the database (VOC2011)
- View the [action guidelines](#) used for annotating the action task images

# Getting started with deep learning

<http://www.deeplearningbook.org/>



## Deep Learning

- [Table of Contents](#)
- [Acknowledgements](#)
- [Notation](#)
- [1 Introduction](#)
- [Part I: Applied Math and Machine Learning Basics](#)
  - [2 Linear Algebra](#)
  - [3 Probability and Information Theory](#)
  - [4 Numerical Computation](#)
  - [5 Machine Learning Basics](#)
- [Part II: Modern Practical Deep Networks](#)
  - [6 Deep Feedforward Networks](#)
  - [7 Regularization for Deep Learning](#)
  - [8 Optimization for Training Deep Models](#)
  - [9 Convolutional Networks](#)
  - [10 Sequence Modeling: Recurrent and Recursive Nets](#)
  - [11 Practical Methodology](#)
  - [12 Applications](#)
- [Part III: Deep Learning Research](#)
  - [13 Linear Factor Models](#)
  - [14 Autoencoders](#)
  - [15 Representation Learning](#)
  - [16 Structured Probabilistic Models for Deep Learning](#)
  - [17 Monte Carlo Methods](#)
  - [18 Confronting the Partition Function](#)
  - [19 Approximate Inference](#)
  - [20 Deep Generative Models](#)

# Landmark papers – worth a read

Imagenet classification with deep convolutional neural networks (AlexNet)

A Krizhevsky, I Sutskever, GE Hinton - Advances in neural ..., 2012

Cited by 10k+

Very deep convolutional networks for large-scale image recognition (VGG)

K Simonyan, A Zisserman – ICLR 2015

Cited by 3k+