Autoencoding beyond pixels using a learned similarity metric

We present an autoencoder that leverages learned representations to better measure similarities in data space. By combining a variational autoencoder (VAE) with a generative adversarial network (GAN) we can use learned feature representations in the GAN discriminator as basis for the VAE reconstruction objective. Thereby, we replace element-wise errors with feature-wise errors to better capture the data distribution while offering invariance towards e.g. translation. We apply our method to images of faces and show that it outperforms VAEs with element-wise similarity measures in terms of visual fidelity. Moreover, we show that the method learns an embedding in which high-level abstract visual features (e.g. wearing glasses) can be modified using simple arithmetic.
DeepPy: Pythonic deep learning
This technical report introduces DeepPy – a deep learning framework built on top of NumPy with GPU acceleration. DeepPy bridges the gap between high performance neural networks and the ease of development from Python/NumPy. Users with a background in scientific computing in Python will quickly be able to understand and change the DeepPy codebase as it is mainly implemented using high-level NumPy primitives. Moreover, DeepPy supports complex network architectures by letting the user compose mathematical expressions as directed graphs. The latest version is available at http://github.com/andersbll/deeppy under the MIT license.

Learned image representations for visual recognition
This thesis addresses the problem of extracting image structures for representing images effectively in order to solve visual recognition tasks. Problems from diverse research areas (medical imaging, material science and food processing) have motivated large parts of the methodological development. The solutions are inspired by and extend state-of-the-art techniques for describing and learning image content.

More specifically, the thesis explores two approaches to constructing image representations, namely feature engineering and feature learning. In the feature engineering approach, we devise a new image representation for texture-like patterns based on count statistics of second-order image structure. We demonstrate the discriminative capabilities of this representation on medical images and perform both cell classification and mitosis detection. Moreover, we develop an object identification method based on vector quantized local image descriptors allowing us to distinguish individual meat cuts along a production line and trace them in a non-intrusive manner. In the feature learning approach, we propose to solve the task of segmenting scanning electron microscopy images of calcite crystals by learning a meaningful pixel description to facilitate the actual segmentation. Finally, we present a new unsupervised generative image model addressing the problem of pixel-based similarity measures for images. We propose a scheme for employing feature-based similarity measures and demonstrate how this improves the ability to learn high-level concepts in images of faces.

The thesis argues in favor of learning features and presents new methods for domains with limited amounts of labeled data allowing feature learning to be applied more broadly.
**Oriented Shape Index Histograms for Cell Classification**

We propose a novel extension to the shape index histogram feature descriptor where the orientation of the second-order curvature is included in the histograms. The orientation of the shape index is reminiscent but not equal to gradient orientation which is widely used for feature description. We evaluate our new feature descriptor using a public dataset consisting of HEp-2 cell images from indirect immunofluorescence lighting. Our results show that we can improve classification performance significantly when including the shape index orientation. Notably, we show that shape index orientation outperforms the gradient orientation on the dataset.

**Assessment of algorithms for mitosis detection in breast cancer histopathology images**

The proliferative activity of breast tumors, which is routinely estimated by counting of mitotic figures in hematoxylin and eosin stained histology sections, is considered to be one of the most important prognostic markers. However, mitosis counting is laborious, subjective and may suffer from low inter-observer agreement. With the wider acceptance of whole slide images in pathology labs, automatic image analysis has been proposed as a potential solution for these issues. In this paper, the results from the Assessment of Mitosis Detection Algorithms 2013 (AMIDA13) challenge are described. The challenge was based on a data set consisting of 12 training and 11 testing subjects, with more than one thousand annotated mitotic figures by multiple observers. Short descriptions and results from the evaluation of eleven methods are presented. The top performing method has an error rate that is comparable to the inter-observer agreement among pathologists.
CUDArray: CUDA-based NumPy

This technical report introduces CUDArray – a CUDA-accelerated subset of the NumPy library. The goal of CUDArray is to combine the ease of development from NumPy with the computational power of Nvidia GPUs in a lightweight and
extensible framework. Since the motivation behind CUDArray is to facilitate neural network programming, CUDArray extends NumPy with a neural network submodule. This module has both a CPU and a GPU back-end to allow for experiments without requiring a GPU.

**HEp-2 Cell Classification Using Shape Index Histograms With Donut-Shaped Spatial Pooling**

We present a new method for automatic classification of indirect immunofluorescence images of HEp-2 cells into different staining pattern classes. Our method is based on a new texture measure called shape index histograms that captures second-order image structure at multiple scales. Moreover, we introduce a spatial decomposition scheme which is radially symmetric and suitable for cell images. The spatial decomposition is performed using donut-shaped pooling regions of varying sizes when gathering histogram contributions. We evaluate our method using both the ICIP 2013 and the ICPR 2012 competition datasets. Our results show that shape index histograms are superior to other popular texture descriptors for HEp-2 cell classification. Moreover, when comparing to other automated systems for HEp-2 cell classification we show that shape index histograms are very competitive; especially considering the relatively low complexity of the method.
Vision-based method for tracking meat cuts in slaughterhouses

Meat traceability is important for linking process and quality parameters from the individual meat cuts back to the production data from the farmer that produced the animal. Current tracking systems rely on physical tagging, which is too intrusive for individual meat cuts in a slaughterhouse environment. In this article, we demonstrate a computer vision system for recognizing meat cuts at different points along a slaughterhouse production line. More specifically, we show that 211 pig loins can be identified correctly between two photo sessions. The pig loins undergo various perturbation scenarios (hanging, rough treatment and incorrect trimming) and our method is able to handle these perturbations gracefully. This study shows that the suggested vision-based approach to tracking is a promising alternative to the more intrusive methods currently available.
An explorative study on pork loin recognition

Bag-of-words (BoW) image description has shown good performance for a large variety of image recognition scenarios. We investigate approaches to alleviating a standard BoW image description pipeline representations for the specific task of recognizing pork loins. Specifically, we extend the BoW description to include depth maps, perform non-rigid image registration to align the images, and apply PCA dimensionality reduction on the BoW descriptors. Our results show that the combination of image registration and PCA yields a more distinctive recognition.

Online Multi-Spectral Meat Inspection

We perform an explorative study on multi-spectral image data from a prototype device developed for fast online quality inspection of meat products. Because the camera setup is built for speed, we sacrifice exact pixel correspondences between the different bands of the multi-spectral images.

Our work is threefold as we 1) investigate the color distributions and construct a model to describe pork loins, 2) classify the different components in pork loins (meat, fat, membrane), and 3) detect foreign objects on the surface of pork loins. Our investigation shows that the color distributions can effectively be modeled using the Gaussian mixture model (GMM). For the classification task we build a classifier using a GMM. For detecting foreign objects, we construct a novelty detector using a GMM.

We evaluate our method on a small dataset with mixed results. While we are able to provide reasonable classifications, the multi-spectral data does not seem to offer significant additional information compared to a standard RGB camera. Moreover, the multi-spectral images come with the cost of losing pixel correspondences.
Pond of Illusion: Interacting through Mixed Reality

Pond of Illusion is a mixed reality installation where a virtual space (the pond) is injected between two real spaces. The users are in either of the real spaces, and they can see each other through windows in the virtual space as illustrated in Figure 1(left). The installation attracts people to a large display in either of the real spaces by allowing them to feed virtual fish swimming in the pond. Figure 1(middle) shows how a Microsoft Kinect mounted on top of the display is used for detecting throw motions, which triggers virtual breadcrumbs to be thrown into the pond for feeding the nearby fish. Of course, the fish may not be available because they are busy eating what people have thrown into the pond from the other side.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Nobel-Jørgensen, M. (Intern), Nielsen, J. B. (Intern), Larsen, A. B. L. (Intern), Olsen, M. D. (Intern), Frisvad, J. R. (Intern), Bærentzen, J. A. (Intern)
Number of pages: 1
Publication date: 2013

Host publication information
Title of host publication: Proceedings of SIGGRAPH Asia 2013 Posters
Publisher: Association for Computing Machinery
Article number: 26
ISBN (Print): 978-1-4503-2634-6
BFI conference series: International Conference on Computer Graphics and Interactive Techniques (5000583)
Main Research Area: Technical/natural sciences
DOIs: 10.1145/2542302.2542334

Bibliographical note
Poster abstract.
Source: dtu
Source-ID: n::oai:DTIC-ART:acm/426516734::34909
Publication: Research - peer-review › Conference abstract in proceedings – Annual report year: 2013

Jet-Based Local Image Descriptors
We present a general novel image descriptor based on higher-order differential geometry and investigate the effect of common descriptor choices. Our investigation is twofold in that we develop a jet-based descriptor and perform a comparative evaluation with current state-of-the-art descriptors on the recently released DTU Robot dataset. We demonstrate how the use of higher-order image structures enables us to reduce the descriptor dimensionality while still achieving very good performance. The descriptors are tested in a variety of scenarios including large changes in scale, viewing angle and lighting. We show that the proposed jet-based descriptor is superior to state-of-the-art for DoG interest points and show competitive performance for the other tested interest points.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics, University of Copenhagen
Authors: Larsen, A. B. L. (Intern), Darkner, S. (Intern), Dahl, A. L. (Intern), Pedersen, K. S. (Forskerdatabase)
Pages: 638-650
Publication date: 2012

Host publication information
Control and Surveillance of Automated Production Steps using Computer Vision

Department of Applied Mathematics and Computer Science
Period: 15/07/2012 → 25/08/2016
Number of participants: 6
PhD Student:
Larsen, Anders Boesen Lindbo (Intern)
Supervisor:
Dahl, Anders Bjorholm (Intern)
Main Supervisor:
Larsen, Rasmus (Intern)
Examiner:
Nielsen, Allan Aasbjerg (Intern)
Belongie, Serge (Ekstern)
Nielsen, Mads (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: Eksternt finansieret virksomhed
Project: PhD

Control & Surveillance of Automated Production Steps (a part of the inSPIRe Food)
Summary of project: Automation of many manual operations in the food industry is difficult, because the criteria for process control are often based on tacit knowledge of the operator. Our hypothesis is that a route to optimal automation of such operations is to register how the trained process operator makes decisions from observations of the process and combining this knowledge with predictive modelling of input/output of the process units.

Department of Applied Mathematics and Computer Science
Statistics and Data Analysis
National Food Institute
Research Group for Food Production Engineering
Image Analysis & Computer Graphics
Period: 01/01/2011 → 31/12/2016
Number of participants: 5
Project participant:
Larsen, Rasmus (Intern)
Ersbøll, Bjarne Kjær (Intern)
Frosch, Stina (Intern)
Clemmensen, Line Katrine Harder (Intern)
Larsen, Anders Boesen Lindbo (Intern)

Financing sources
Source: Public research council
Name of research programme: Danish Council for Strategic Research and the Danish Council for Technology (now The Danish Innovation Foundation)
Amount: 5,218,000.00 Danish Kroner
Project