Reference List

Tyler (Taewook) Kim

June 20, 2019

Contents

1 Deep Neural Networks

1.1 Deep Learning Book

(7) - *Deep Learning Book*

Abstract:

The text offers mathematical and conceptual background, covering relevant concepts in linear algebra, probability theory and information theory, numerical computation, and machine learning. It describes deep learning techniques used by practitioners in industry, including deep feedforward networks, regularization, optimization algorithms, convolutional networks, sequence modeling, and practical methodology; and it surveys such applications as natural language processing, speech recognition, computer vision, online recommendation systems, bioinformatics, and videogames. Finally, the book offers research perspectives, covering such theoretical topics as linear factor models, autoencoders, representation learning, structured probabilistic models, Monte Carlo methods, the partition function, approximate inference, and deep generative models.
Notes: ch. 3, 5, 9, 12 + SGD, Curse of Dimensionality, MLP.

1.2 Deep Convolutional Neural Networks

(6) - *Deep Convolutional Neural Networks*

Abstract:

Neural networks are a subset of the field of artificial intelligence (AI). The predominant types of neural networks used for multidimensional signal processing are deep convolutional neural networks (CNNs). The term deep refers generically to networks having from a "few" to several dozen or more convolution layers, and deep learning refers to methodologies for training these systems to automatically learn their functional parameters using data representative of a specific problem domain of interest. CNNs are currently being used in a broad spectrum of application areas, all of which share the common objective of being able to automatically learn features from (typically massive) data bases and to generalize their responses to circumstances not encountered during the learning phase. Ultimately, the learned features can be used for tasks such as classifying the types of signals the CNN is expected to process. The purpose of this "Lecture Notes" article is twofold: 1) to introduce the fundamental architecture of CNNs and 2) to illustrate, via a computational example, how CNNs are trained and used in practice to solve a specific class of problems..
Notes: Introduces the fundamental architecture of CNN.

2 Existing Approaches to Texture Analysis

2.1 A Review of Texture Classification Methods and Databases

(2) - *A Review of Texture Classification Methods and Databases*

Abstract:

In this survey, we present a review of methods and resources for texture recognition, presenting the most common techniques that have been used in the recent decades, along with current tendencies. That said, this paper covers since the most traditional approaches, for instance texture descriptors such as gray-level co-occurence matrices (GLCM) and Local Binary Patterns (LBP), to more recent approaches such as Convolutional Neural Networks (CNN) and multi-scale patch-based recognition based on encoding approaches such as Fisher Vectors. In addi- tion, we point out relevant references for benchmark datasets, which can help the reader develop and evaluate new methods.

Notes: Discusses various texture analysis methods such as LBP and GLCM.

2.2 Deep TEN: Texture Encoding Network

(13) - *Deep TEN: Texture Encoding Network*

Abstract:

We propose a Deep Texture Encoding Network (Deep-TEN) with a novel Encoding Layer integrated on top of convolutional layers, which ports the entire dictionary learn- ing and encoding pipeline into a single model. Current methods build from distinct components, using standard encoders with separate off-the-shelf features such as SIFT descriptors or pre-trained CNN features for material recognition. Our new approach provides an end-to-end learning framework, where the inherent visual vocabularies are learned directly from the loss function. The features, dictionaries and the encoding representation for the classifier are all learned simultaneously. The representation is orderless and therefore is particularly useful for material and texture recognition. The Encoding Layer generalizes robust residual encoders such as VLAD and Fisher Vec- tors, and has the property of discarding domain specific information which makes the learned convolutional features easier to transfer. Additionally, joint training using multiple datasets of varied sizes and class labels is supported resulting in increased recognition performance. The experimental results show superior performance as com- pared to state-of-the-art methods using gold-standard databases such as MINC-2500, Flickr Material Database, KTH-TIPS-2b, and two recent databases 4D-Light-Field- Material and GTOS. The source code for the complete system are publicly available.

Notes: Integrated encoding layer on top of convolutional layers.

2.3 CNN based on LBP for Evaluating Natural Disasters

(3) - *CNN based on LBP for Evaluating Natural Disasters*

Abstract:

This paper presents a novel evaluation method of areas affected by natural disasters with the purpose of managing these crisis situations. Since it is necessary to have a real overview of a specific area in the shortest time, our methodology proposes a neural network with backpropagation approach for flood detection from UAV images. For this, the Local Binary Pattern (LBP) texture operator is used for areas classification. The LBP operator labels each pixel of the analyzed image by comparing it with its neighbors, which ends with the computation of a binary number that it is converted to decimal format named LBP code. Thus, based on the generated LBP codes, a histogram type feature is computed and used in both training and testing phases of the proposed neural network. Over 50 images obtained with the aid of UAV technology were tested with the proposed neural network and good results in terms of accuracy for flood areas detection were obtained.

Notes: looks REALLY interesting; computed histogram type feature based on the generated LBP codes

3 Hand-Crafted Features

3.1 Histograms of oriented gradients for human detection

(4) - *Histograms of oriented gradients for human detection*

Abstract:

We study the question of feature sets for robust visual object recognition; adopting linear SVM based human detection as a test case. After reviewing existing edge and gradient based descriptors, we show experimentally that grids of histograms of oriented gradient (HOG) descriptors significantly outperform existing feature sets for human detection. We study the influence of each stage of the computation on performance, concluding that fine-scale gradients, fine orientation binning, relatively coarse spatial binning, and high-quality local contrast normalization in overlapping descriptor blocks are all important for good results. The new approach gives near-perfect separation on the original MIT pedestrian database, so we introduce a more challenging dataset containing over 1800 annotated human images with a large range of pose variations and backgrounds.
Notes: Slide3 - 1

3.2 Multiresolution gray-scale and rotation invariant texture classification with local binary patterns

(11) - *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*

Abstract:

Presents a theoretically very simple, yet efficient, multiresolution approach to gray-scale and rotation invariant texture classification based on local binary patterns and nonparametric discrimination of sample and prototype distributions. The method is based on recognizing that certain local binary patterns, termed "uniform," are fundamental properties of local image texture and their occurrence histogram is proven to be a very powerful texture feature. We derive a generalized gray-scale and rotation invariant operator presentation that allows for detecting the "uniform" patterns for any quantization of the angular space and for any spatial resolution and presents a method for combining multiple operators for multiresolution analysis. The proposed approach is very robust in terms of gray-scale variations since the operator is, by definition, invariant against any monotonic transformation of the gray scale. Another advantage is computational simplicity as the operator can be realized with a few operations in a small neighborhood and a lookup table. Experimental results demonstrate that good discrimination can be achieved with the occurrence statistics of simple rotation invariant local binary patterns.
Notes: Slide3 - 2

3.3 Detection and Discrimination of Land Mines in Ground-Penetrating Radar Based on Edge Histogram Descriptors and a Possibilistic K-Nearest Neighbor Classifier

(5) - *Detection and Discrimination of Land Mines in Ground-Penetrating Radar Based on Edge Histogram Descriptors and a Possibilistic K-Nearest Neighbor Classifier*

Abstract:

This paper describes an algorithm for land mine detection using sensor data generated by a ground-penetrating radar (GPR) system that uses edge histogram descriptors for feature extraction and a possibilistic K -nearest neighbors (K -NNs) rule for confidence assignment. The algorithm demonstrated the best performance among several high-performance algorithms in extensive testing on a large real-world datasets associated with the difficult problem of land mine detection. The superior performance of the algorithm is attributed to the use of the possibilistic K -NN algorithm, thereby providing important evidence supporting the use of possibilistic methods in real-world applications. The GPR produces a 3-D array of intensity values, representing a volume below the surface of the ground. First, a computationally inexpensive prescreening algorithm for anomaly detection is used to focus attention and identify candidate signatures that resemble mines. The identified regions of interest are processed further by a feature extraction algorithm to capture their salient features. We use translation-invariant features that are based on the local edge distribution of the 3-D GPR signatures. Specifically, each 3-D signature is divided into subsignatures, and the local edge distribution for each subsignature is represented by a histogram. Next, the training signatures are clustered to identify prototypes. The main idea is to identify few prototypes that can capture the variations of the signatures within each class. These variations could be due to different mine types, different soil conditions, different weather conditions, etc. Fuzzy memberships are assigned to these representatives to capture their degree of sharing among the mines and false alarm classes. Finally, a possibilistic K -NN-based rule is used to assign a confidence value to distinguish true detections from false alarms. The proposed algorithm is implemented and integrated within a complete land mine prototype system. It is trained, field-tested, evaluated, and compared using a large-scale cross-validation experiment that uses a diverse dataset acquired from four outdoor test sites at different geographic locations. This collection covers over 41 807 m 2 of ground and includes 1593 mine encounters.

Notes: Slide3 - 3

4 Motivation for Histogram Layer

4.1 Augmenting a convolutional neural network with local histograms - A case study in crop classification from high-resolution UAV imagery

(12) - *Augmenting a convolutional neural network with local histograms - A case study in crop classification from high-resolution UAV imagery*

Abstract:

The advent of affordable drones capable of taking high resolution images of agricultural fields creates new challenges and opportunities in aerial scene understanding. This paper tackles the problem of recognizing crop types from aerial imagery and proposes a new hybrid neural network architecture which combines histograms and convolutional units. We evaluate the performance of the hybrid model on a 23-class classification task and compare it to convolutional and histogram-based models. The result is an improvement of the classification performance.

Notes: Very similar to what I wrote in my application essay; mentions HistCNN

4.2 Deep neural networks for texture classification—A theoretical analysis

(1) - *Deep neural networks for texture classification—A theoretical analysis*

Abstract:

We investigate the use of Deep Neural Networks for the classification of image datasets where texture features are important for generating class-conditional discriminative representations. To this end, we first derive the size of the feature space for some standard textural features extracted from the input dataset and then use the theory of Vapnik–Chervonenkis dimension to show that hand-crafted feature extraction creates low-dimensional representations which help in reducing the overall excess error rate. As a corollary to this analysis, we derive for the first time upper bounds on the VC dimension of Convolutional Neural Network as well as Dropout and Dropconnect networks and the relation between excess error rate of Dropout and Dropconnect networks. The concept of intrinsic dimension is used to validate the intuition that texture-based datasets are inherently higher dimensional as compared to handwritten digits or other object recognition datasets and hence more difficult to be shattered by neural networks. We then derive the mean distance from the centroid to the nearest and farthest sampling points in an n-dimensional manifold and show that the Relative Contrast of the sample data vanishes as dimensionality of the underlying vector space tends to infinity.

Notes: Deep learning + Texture Analysis.

5 Existing Hybrid Models of Deep Learning and Hand-Crafted Features

5.1 A hybrid of deep learning and hand-crafted features based approach for snow cover mapping

(10) - *A hybrid of deep learning and hand-crafted features based approach for snow cover mapping*

Abstract:

Monitoring the extent of snow cover plays a vital role for a better understanding of current and future climatic, ecological, and water cycle conditions. Previously, several traditional machine learning models have been applied for accomplishing this while exploring a variety of feature extraction techniques on various information sources. However, the laborious process of any amount of hand-crafted feature extraction has not helped to obtain high accuracies. Recently, deep learning models have shown that feature extraction can be made automatic and that they can achieve the required high accuracies but at the cost of requiring a large amount of labelled data. Fortunately, despite the absence of such large amounts of labelled data for this task, we can rely on pre-trained models, which accept red-green-blue (RGB) information (or dimensions-reduced spectral data). However, it is always better to include a variety of information sources to solve any problem, especially with the availability of other important information sources like synthetic aperture radar (SAR) imagery and elevation. We propose a hybrid model where the deep learning is assisted by these information sources which have until now been left out. Particularly, our model learns from both the deep learning features (derived from spectral data) and the hand-crafted features (derived from SAR and elevation). Such an approach shows interesting performance-improvement from 96.02

Notes: Discusses pro/con for DL and HCF.

5.2 From BoW to CNN: Two Decades of Texture Representation for Texture Classification

(8) - *From BoW to CNN: Two Decades of Texture Representation for Texture Classification*

Abstract:

Texture is a fundamental characteristic of many types of images, and texture representation is one of the essential and challenging problems in computer vision and pattern recognition which has attracted extensive research attention over several decades. Since 2000, texture representations based on Bag of Words and on Convolutional Neural Networks have been extensively studied with impressive performance. Given this period of remarkable evolution, this paper aims to present a comprehensive survey of advances in texture representation over the last two decades. More than 250 major publications are cited in this survey covering different aspects of the research, including benchmark datasets and state of the art results. In retrospect of what has been achieved so far, the survey discusses open challenges and directions for future research.

Notes: Discusses pro/con for DL and HCF.

5.3 Combining Deep and Handcrafted Image Features for Presentation Attack Detection in Face Recognition Systems Using Visible-Light Camera Sensors

(9) - *Combining Deep and Handcrafted Image Features for Presentation Attack Detection in Face Recognition Systems Using Visible-Light Camera Sensors*

Abstract:

Although face recognition systems have wide application, they are vulnerable to presentation attack samples (fake samples). Therefore, a presentation attack detection (PAD) method is required to enhance the security level of face recognition systems. Most of the previously proposed PAD methods for face recognition systems have focused on using handcrafted image features, which are designed by expert knowledge of designers, such as Gabor filter, local binary pattern (LBP), local ternary pattern (LTP), and histogram of oriented gradients (HOG). As a result, the extracted features reflect limited aspects of the problem, yielding a detection accuracy that is low and varies with the characteristics of presentation attack face images. The deep learning method has been developed in the computer vision research community, which is proven to be suitable for automatically training a feature extractor that can be used to enhance the ability of handcrafted features. To overcome the limitations of previously proposed PAD methods, we propose a new PAD method that uses a combination of deep and handcrafted features extracted from the images by visible-light camera sensor. Our proposed method uses the convolutional neural network (CNN) method to extract deep image features and the multi-level local binary pattern (MLBP) method to extract skin detail features from face images to discriminate the real and presentation attack face images. By combining the two types of image features, we form a new type of image features, called hybrid features, which has stronger discrimination ability than single image features. Finally, we use the support vector machine (SVM) method to classify the image features into real or presentation attack class. Our experimental results indicate that our proposed method outperforms previous PAD methods by yielding the smallest error rates on the same image databases.

Notes: Mentions LBP HOG; perhaps extension of texture analysis?

References

[1] Saikat Basu, Supratik Mukhopadhyay, Manohar Karki, Robert DiBiano, Sangram Ganguly, Ramakrishna Nemani, and Shreekant Gayaka. Deep neural networks for texture classification—a theoretical analysis. *Neural Networks*, 97:173 – 182, 2018. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2017.10.001. URL http://www.sciencedirect.com/science/article/pii/S0893608017302320.

[2] P. Cavalin and L. S. Oliveira. A review of texture classification methods and databases. In *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T)*, pages 1–8, Oct 2017. doi: 10.1109/SIBGRAPI-T.2017.10.

[3] Andrada Livia Cirneanu, Dan Popescu, and Loretta Ichim. Cnn based on lbp for evaluating natural disasters. *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pages 568–573, 2018.

[4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1, June 2005. doi: 10.1109/CVPR.2005.177.

[5] H. Frigui and P. Gader. Detection and discrimination of land mines in ground-penetrating radar based on edge histogram descriptors and a possibilistic k-nearest neighbor classifier. *IEEE Transactions on Fuzzy Systems*, 17(1):185–199, Feb 2009. ISSN 1063-6706. doi: 10.1109/TFUZZ.2008.2005249.

[6] R. C. Gonzalez. Deep convolutional neural networks [lecture notes]. *IEEE Signal Processing Magazine*, 35(6):79–87, Nov 2018. ISSN 1053-5888. doi: 10.1109/MSP.2018.2842646.

[7] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[8] Li Liu, Jie Chen, Paul Fieguth, Guoying Zhao, Rama Chellappa, and Matti Pietikäinen. From bow to cnn: Two decades of texture representation for texture classification. *International Journal of Computer Vision*, 127(1):74–109, Jan 2019. ISSN 1573-1405. doi: 10.1007/s11263-018-1125-z. URL https://doi.org/10.1007/s11263-018-1125-z.

[9] Dat Nguyen, Tuyen Pham, Na Rae Baek, and Kang Ryoung Park. Combining deep and handcrafted image features for presentation attack detection in face recognition systems using visible-light camera sensors. *Sensors*, 18:699, 02 2018. doi: 10.3390/s18030699.

[10] Rahul Nijhawan, Josodhir Das, and Balasubramanian Raman. A hybrid of deep learning and handcrafted features based approach for snow cover mapping. *International Journal of Remote Sensing*, 40(2):759–773, 2019. doi: 10.1080/01431161.2018.1519277. URL https://doi.org/10.1080/01431161.2018.1519277.

[11] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, July 2002. ISSN 0162-8828. doi: 10.1109/TPAMI.2002.1017623.

[12] Julien Rebetez, Héctor F Satizábal, Matteo Mota, Dorothea Noll, Lucie Büchi, Marina Wendling, Bertrand Cannelle, Andrés Pérez-Uribe, and Stéphane Burgos. Augmenting a convolutional neural network with local histograms-a case study in crop classification from high-resolution uav imagery. In *ESANN*, 2016.

[13] Hang Zhang, Jia Xue, and Kristin J. Dana. Deep TEN: texture encoding network. *CoRR*, abs/1612.02844, 2016. URL http://arxiv.org/abs/1612.02844.