Master thesis topic (Masterproefonderwerp)	Convolutional Neural Networks (CNNs) analysed via sparse coding
Promoters (Promotoren)	prof. dr. ir. Aleksandra Pižurica – Department of Telecommunications and Information Processing
Mentors (Begeleiders)	Srđan Lazendić – Department of Electronics and Information Systems Shaoguang Huang – Department of Telecomunications and Information Processing
Contact person (Contactpersoon)	Srđan Lazendić
Number of master theses (Aantal masterproeven)	1
Number of students (Aantal studenten)	1
Academic year (Academiejaar)	2020-2021
Keywords (Trefwoorden)	Image processing, Convolutional neural networks, Dictionary learning, Sparse coding
Location – (Locatie)	Technicum, De Sterre, thuis

References (Referenties):

1. W. Samek, T. Wiegand, K.R. Müller, *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models*, 2018.
2. P. Vardan, Y. Romano, M. Elad. *Convolutional neural networks analyzed via convolutional sparse coding*, 2017.
3. V. Papyan, Y. Romano, J. Sulam, M. Elad, *Theoretical Foundations of Deep Learning via Sparse Representations*, 2018.

Description of the problem (Probleemstelling)	Recently, the use of deep networks has led to unprecedented results across various fields of image and data processing. Although these models reach impressive prediction accuracies, due to their non-linear structure, it is not clear what information in the input data makes them actually arrive at their decisions. The configuration and training of deep learning networks are largely driven by trial-and-error strategies. Since this lack of transparency can be a major drawback, the development of methods for explaining and interpreting deep learning models has recently attracted increasing attention. Deep learning as an instance of general representation learning is naturally connected to sparse signal representations and dictionary learning. While the development of new variants of deep neural networks have been largely driven by a considerable amount of intuition, dictionary learning offers a sound theoretical formulation. Hence, there is a lot of interest in combining the two into a powerful, yet better interpretable framework. Many problems like image and video restoration benefit from the sparse representation model. The goal of such a sparse representation model is to approximate well a signal using only few elements from a (typically redundant) dictionary. Recently, the convolutional sparse coding (CSC) paradigm has been introduced. CSC is a special case of the sparse representation model, built around a very structured dictionary being a union of banded and circulant matrices (see Fig. 1). An extension of the CSC model, known as multi-layer sparse model has raised insightful connections between sparse representations and convolutional neural networks (CNN) (see Fig. 2). The multi-layer CSC leads to a solid and systematic theoretical justification of the architectures used in deep learning for CNN networks. However, many aspects of this approach are yet to be explored, both theoretically and in terms of the practical design. Taking into account that this approach allows analysis of CNNs architectures and suggests how to build new ones in a systematic fashion, different sparse coding techniques should be analyzed and compared since different architectures are resulting from different solvers for the features. Student will be guided by the supervisors from the research group GAIM whose research expertise is on representation learning, deep learning and sparse coding.

Goal of the thesis (Doelstelling)	The goal of the thesis is to build on recent works in multi-scale convolutional sparse coding. Firstly, the student should study and understand the theory behind the representation learning, deep learning and dictionary learning in particular. Secondly, concrete solvers for the features (e.g. ML-ISTA, ML-FISTA, LBP, ML-LISTA...) should be chosen and the performance of the constructed architectures should be compared both among themselves and against some of the more conventional tools for image processing tasks. Practical applications will be chosen in the agreement with the student based on his/her affinities. Possible applications are in large-scale hyperspectral data processing in remote sensing and multimodal data analysis in art investigation. The existing code and literature will be made available to the student.

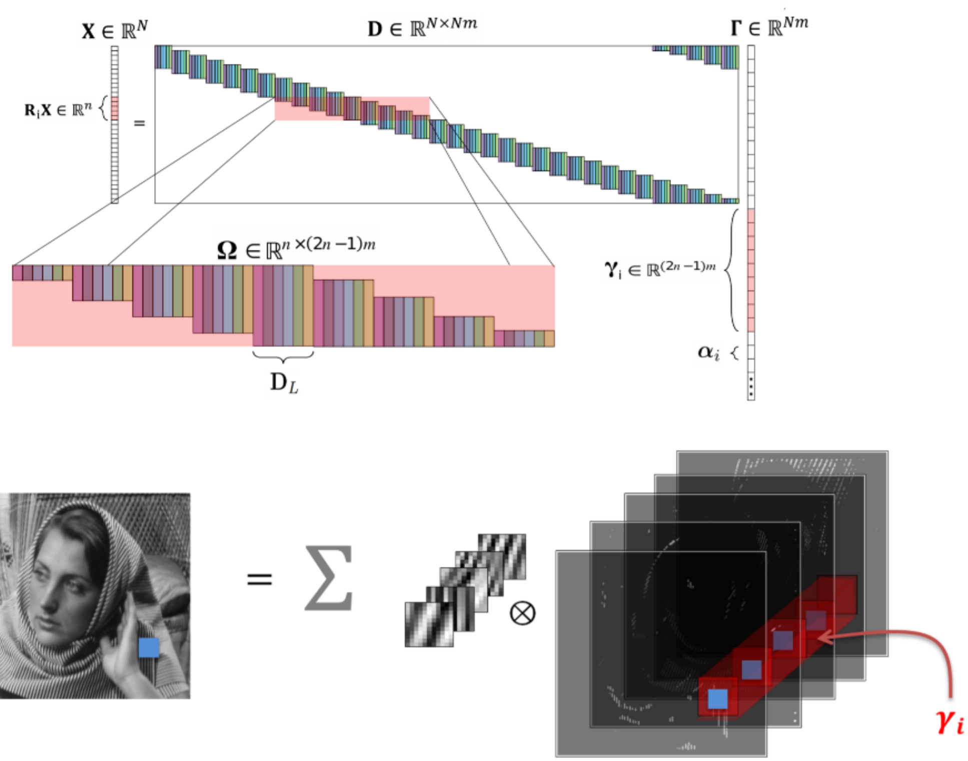

Figure 1: The convolutional model description with the composition in terms of the local dictionary (see [1]).

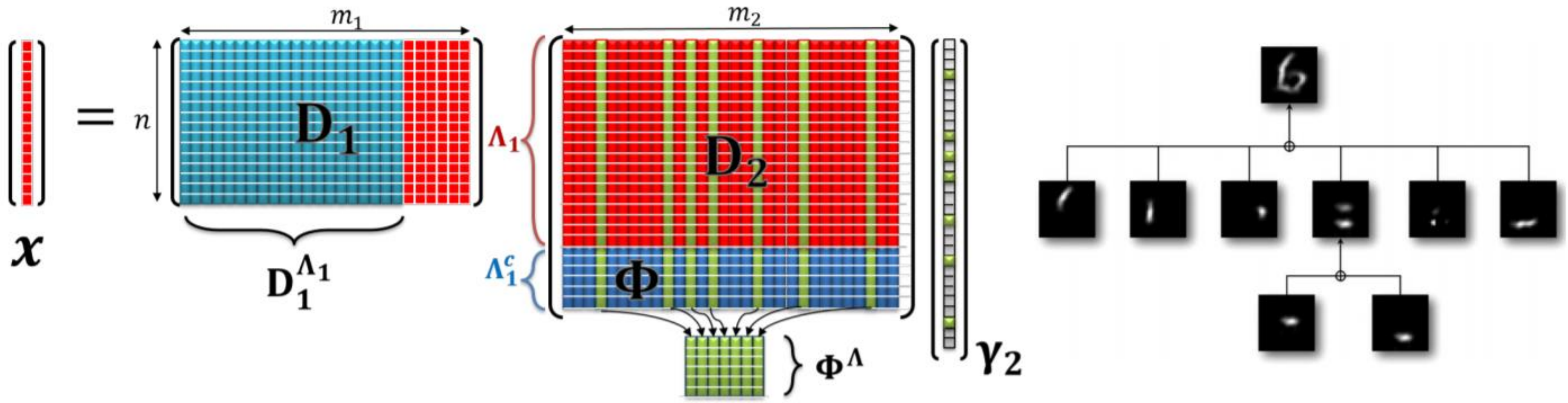

Figure 2: Decomposition of an image from MNIST in terms of two multilayer convolutional dictionaries. Two local convolutional atoms (bottom row) are combined to create molecules – at the second level, which are then combined to create the global atom (number 6) (see [2]).