MACHINE LEARNING FOR HEALTH

Medical Computer Vision: Beyond the Human Eye

CNNs, classic architectures, transfer learning, and validation in medical imaging

TECHNICAL HANDOUT · SESSION 07

Professor Flávio Luiz Seixas
Graduate Program in Computing

Academic teaching material based exclusively on the PDFs available in the project

Table of Contents

Introduction

Medical images combine high dimensionality, spatial structure, and acquisition metadata. Vision models must learn relevant patterns without confusing anatomy and disease with features of the equipment, institution, or protocol. [1]

This session presents how convolutional networks work, distinguishes classification, detection, and segmentation tasks, and discusses transfer learning, data augmentation, and validation by patient and institution. [1]

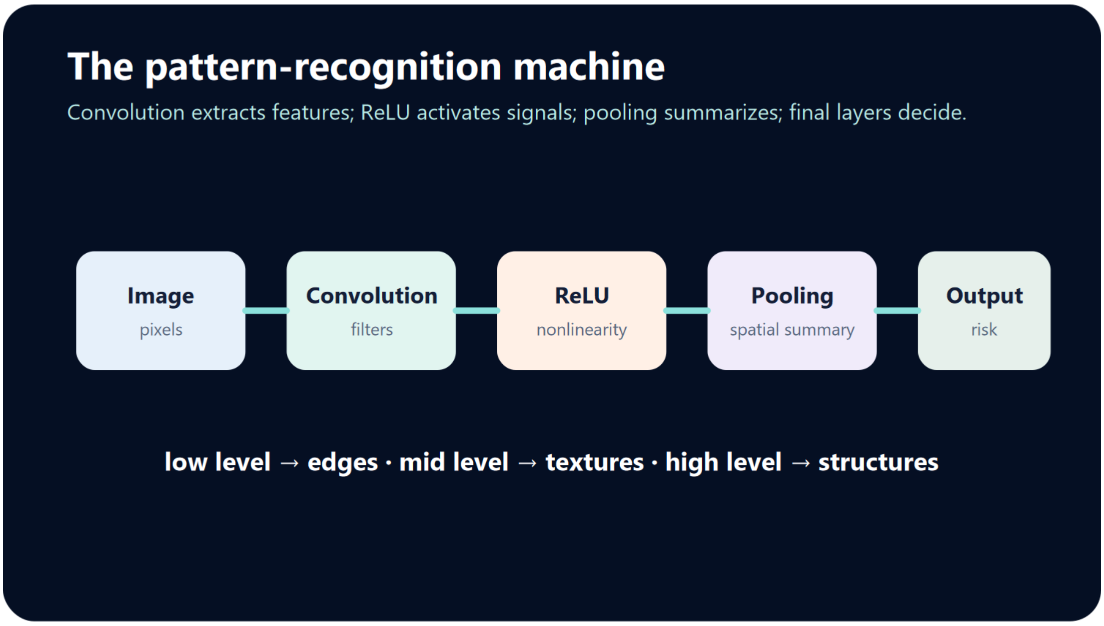

Figure 1 — Building blocks of a convolutional network. Source: original illustration based on the course material and [1].

1. From image to tensor

A digital image represents intensities on a grid and may contain channels or volumes. Medical modalities differ in acquisition mechanism, resolution, contrast, and artifacts. DICOM supports exchange and associates pixels with information needed to interpret and manage the study [1].

KEY CONCEPT
A digital image represents intensities on a grid and may contain channels or volumes.

2. Convolution and feature hierarchies

A convolutional layer applies local filters to the input tensor and produces feature maps. Successive layers can combine simple patterns into more abstract representations; pooling reduces resolution and aggregates responses. Learning adjusts weights from the examples and loss function [1].

Classification produces a label or probability for an image; detection locates objects by region; segmentation assigns a class to pixels or voxels. These outputs require different references and metrics [1].

3. Architectures and transfer learning

Deep architectures vary in how layers and information paths are organized. Residual networks facilitate the training of deeper networks through shortcut connections. U-Net combines an encoding path with a decoding path and connections across scales, making it suitable for segmentation [2].

In transfer learning, weights obtained from another dataset initialize the network and are adjusted for the current task. This strategy can help when few examples are available, but differences between source images and the medical domain must be evaluated [1].

4. Data augmentation and validity

Data augmentation generates variations of examples during training. Transformations are valid only when they preserve clinical meaning. Splits must also be made by patient and, when necessary, by institution to prevent repetition and measure dependence on provenance [3].

Performance should be interpreted together with label quality, population composition, equipment, and protocol. Shortcuts can sustain high local accuracy and fail when the context changes [3].

Session summary

From image to tensor: a digital image represents intensities on a grid and may contain channels or volumes.
Convolution and feature hierarchies: a convolutional layer applies local filters to the input tensor and produces feature maps.
Architectures and transfer learning: deep architectures vary in how layers and information paths are organized.
Data augmentation and validity: data augmentation generates variations of examples during training.

References

[1] Rubin, D. L.; Greenspan, H.; Hoogi, A. Biomedical Imaging Informatics. In: Biomedical Informatics. Springer, 2021. External link

Local file: pdfs/2-Biomedical Informatics/X-Biomedical Imaging Informatics.pdf · PDF pages used: 2, 16, 26, 33, 34, 38, 40.

[2] Kahn, C. E. et al. Interpreting Medical Images. In: Intelligent Systems in Medicine and Health. Springer, 2022. External link

Local file: pdfs/3-Intelligent Systems/XII-Interpreting Medical Images.pdf · PDF pages used: 1, 3, 10, 18.

[3] Subramanian, D.; Cohen, T. A. Machine Learning Systems. In: Intelligent Systems in Medicine and Health. Springer, 2022. External link

Local file: pdfs/3-Intelligent Systems/VI-Machine Learning Systems.pdf · PDF pages used: 20, 22, 52, 57.