Universiteit Leiden
The Netherlands

Deep learning solutions for domain-specific image segmentation

Vadineanu, S.

Summary

Image segmentation is a process that divides an image into distinct regions, identifying and categorising them based on shared characteristics such as colour, texture, or boundaries (see Figure S1 for a schematic representation of cell segmentation). This process has the potential to address a wide range of problems in specialised fields, such as detecting tumours in computed tomography scans for medical imaging or identifying sites in archaeological research. Currently, the best-performing image segmentation algorithms are based on deep learning.

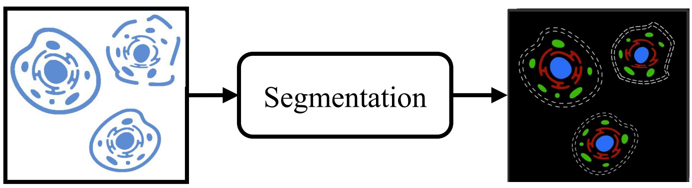

Figure S1: Example of image segmentation. The cell shapes and the segmentation are from [86].

Deep learning refers to a category of statistical models trained to perform tasks by learning from large datasets. For example, in image segmentation, the model is trained using pairs of input images and their corresponding annotations (i.e., categorised regions within the image). Consequently, a learning pipeline for deep learning segmentation algorithms typically includes an initial step in which annotations are generated to prepare for training (annotation process) and a second step where the algorithm is trained using the input images and the previously created annotations (training process). Figure S2 shows an overview of this pipeline.

Despite the promising results of deep learning in image segmentation, its widespread adoption in specialised domains remains hindered by challenges related to annotation

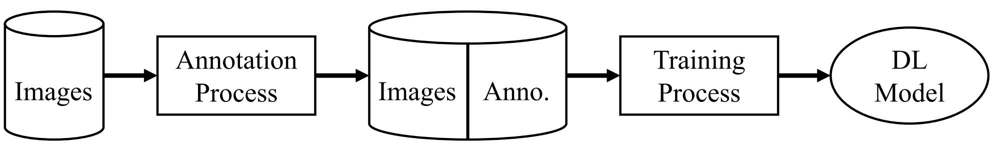

Figure S2: Deep learning pipeline from an initial set of images to a trained model.

and training. When it comes to annotation, marking the regions of an image is not only time-consuming—since it's often done manually—but also prone to mistakes due to human error or unclear boundaries. In specialised domains, this becomes even more difficult because annotators need expert knowledge, which adds costs in terms of time and availability. Experts in these fields often have demanding schedules, and the necessary knowledge may only be possessed by a small group of people. On the training side, deep learning algorithms need large amounts of data and operate with an opaque decision-making process, posing additional barriers. Collecting enough data can be expensive, requiring special equipment, preparing samples (as in medical imaging), or even travelling to specific locations (as in archaeology). Additionally, because deep learning models don't provide a clear, step-by-step explanation of their decisions, professionals in non-technical fields may be hesitant to rely on them.

This thesis provides a set of solutions to address the challenges associated with the annotation and training processes of deep learning algorithms for image segmentation in specialised domains. Specifically, we focus on two applications: cell segmentation in biomedical imaging and the detection of archaeological sites from satellite images. In Chapter 2, we investigate the impact of annotation errors on the performance of deep learning models for cell segmentation. Building on these findings, Chapter 3 introduces a training technique that enables deep learning models to learn from low-quality annotations, such as those with missing regions or imprecise boundaries. In Chapter 4, we propose a novel deep learning algorithm capable of performing cell segmentation using only a few annotated image-annotation pairs (e.g., 5 pairs). Chapter 5 explores the application of explainability techniques to deep learning models trained for image classification, generating visual representations of the image regions the model considers relevant. We analyse these visualisations to gain insights into the learned patterns and further refine them to create semi-automated annotations for archaeological site segmentation, reducing the time required from domain experts compared to manual annotation.