Deep Learning Models in Medical Images

Willy Rodriguez

Torus Actions

October 1, 2024

Outline

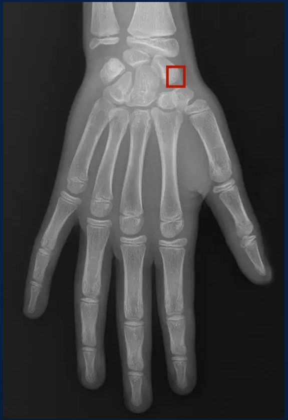
Introduction

Torus Actions. French startup founded in 2019

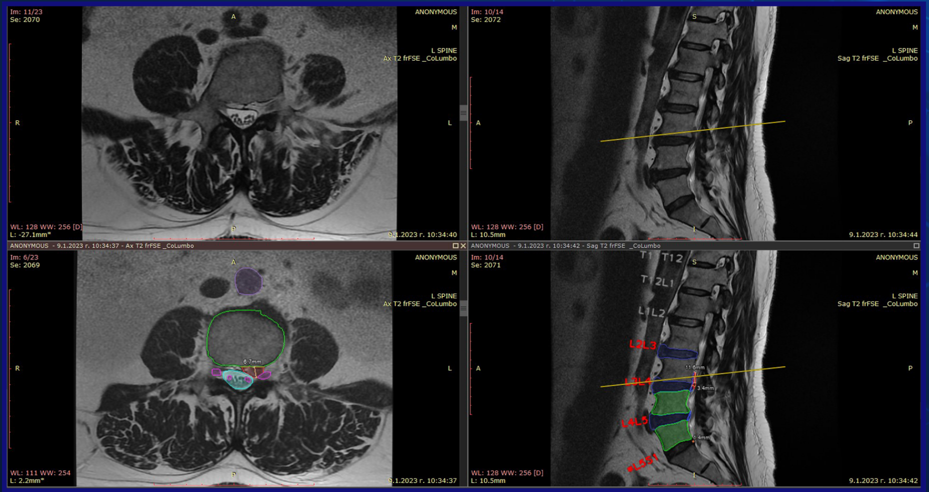
Introduction

Torus AI

Medical Images analysis by AI: Fracture detection

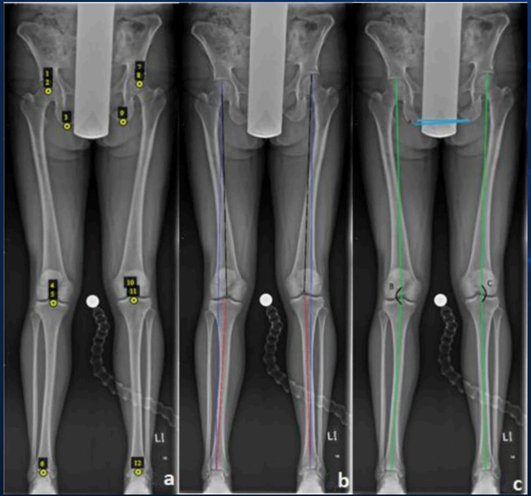


Medical Images analysis by AI: Region identification



Introduction

Medical Images analysis by AI: Automatic measurements

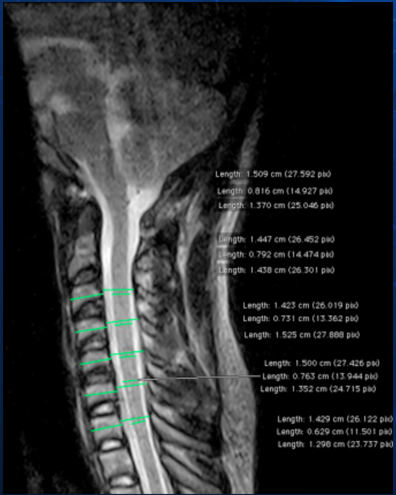


Collaboration with Toulouse Hospital. Cervical Spine and Risk of Severe Lesions

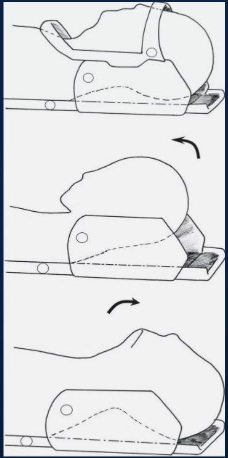
Parcours européen du Stade toulousain

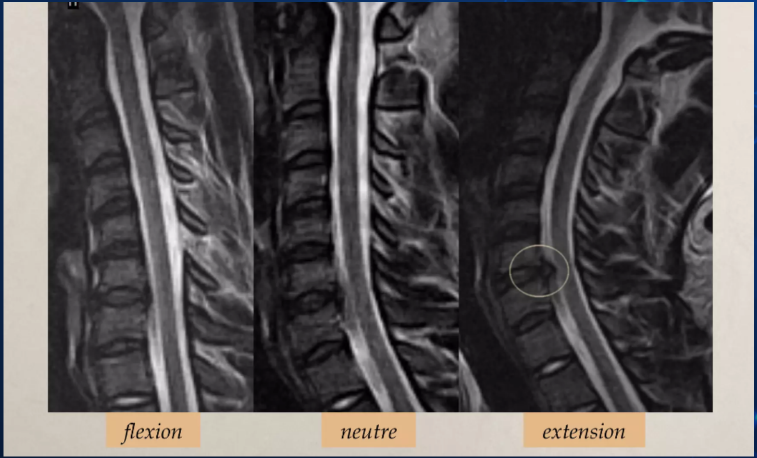
Palmarès et résultats notables	
Titre	6 Coupes d'Europe en 1996, 2003, 2005, 2010, 2021 et 2024
Finale	Coupes d'Europe 2004 et 2008
Demi-finale	Coupes d'Europe 1997, 1998, 2000, 2011, 2019, 2020, 2022 et 2023
Quart de finale	Coupes d'Europe 1999, 2006, 2009, 2012, 2014 et 2017 Challenge européen 2013
Matchs joués	184 (125V, 5N, 54D)
Plus large victoire	Ebbw Vale RFC (108-16) Coupe d'Europe 1998-1999
Plus large défaite	London Wasps (77-17) Coupe d'Europe 1996-1997
Joueur le plus capé	Clément Poitrenaud (96)
Meilleur marqueur	Vincent Clerc (36 essais)
Meilleur réalisateur	Jean-Baptiste Élissalde (437 points)

Collaboration with Toulouse Hospital. Calculation of the medulla-to-cervical canal ratio



Collaboration with Toulouse Hospital. Calculation of the medulla-to-cervical canal ratio (MCR)



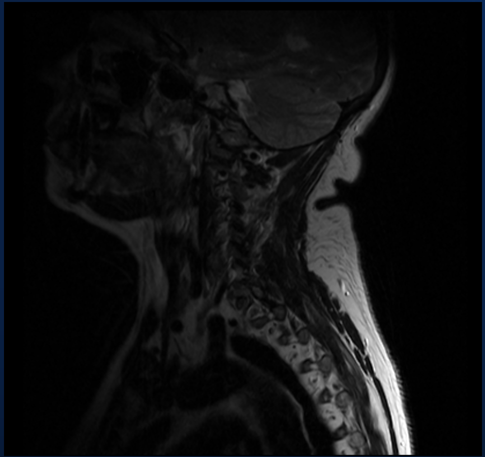


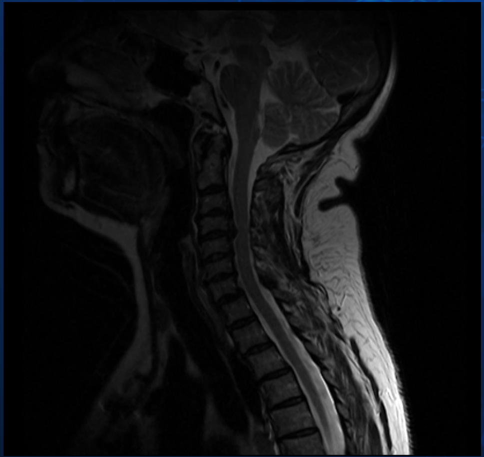
Designing an AI-powered solution for Cervical Spine Analysis

Designing an AI-powered solution for Cervical Spine Analysis

Torus AI

The Data: MRI of cervical spine.





Algorithm outline

1. **Identify the exploitable images and choose the best one**
   - Analyze the images in the dicom series.
   - Select the image with the highest quality for further processing.
   - Technique: VAE for unsupervised learning.

2. **Detect the Regions of Interest (ROI)**
   - Apply image processing techniques to identify important anatomical structures.
   - Technique: Segmentation model using a U-Net architecture.

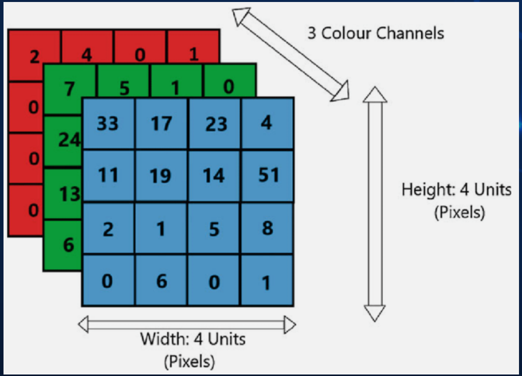
3. **Post-process the results and perform computations**
   - Identify the keypoints for computations.
   - Execute the necessary calculations (distance and ratios).
   - Technique: Geometry and some boring post-processing.

What is an RGB Image?

- An image is made up of tiny squares called **pixels**.
- In a digital image, each pixel has a color defined by three components:
  - **R**: Red
  - **G**: Green
  - **B**: Blue
- Each component (R, G, B) is an intensity value ranging from 0 to 255.

RGB Image as a 3D Tensor

- A 256x256 RGB image is represented as a **3x256x256** tensor:
- **3**: Number of color channels (Red, Green, Blue).
- **256x256**: Width and height of the image (256 pixels wide, 256 pixels tall).

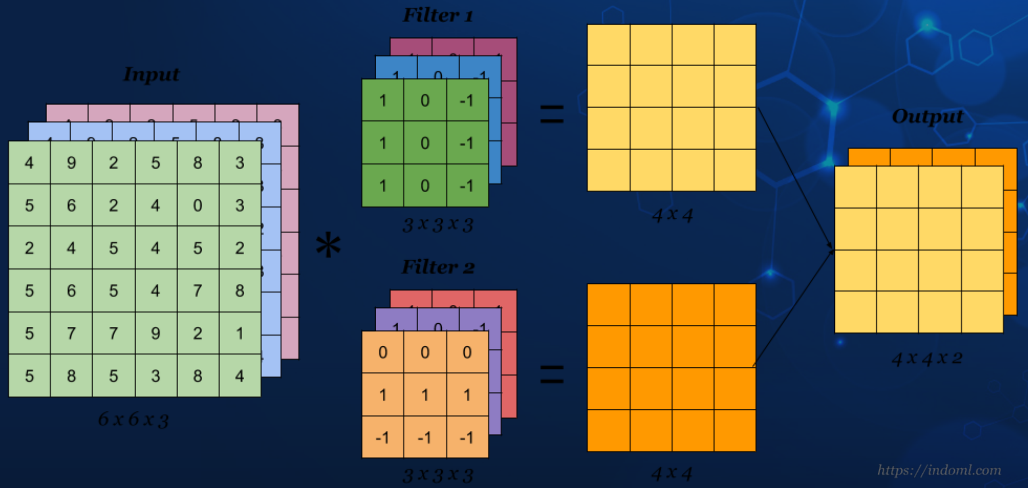


Basic notions

- **Neural Network**: function composed by the concatenation of multiple linear blocks followed by activation functions.
- **Layers**:
  - **Dense Layer**: Linear function
  - **Convolution Layer**: Apply a convolution operation on the input.
  - **Concatenation Layer**: It's possible to compose multiple layers (input, hidden, output) combined with activation functions.
- **Activation Function**:
  - Placed between two linear operations. Ensure the non-linearity of the Neural Network (e.g., **ReLU**, **Sigmoid**).
- **Training the Network**:
  - The network adjusts its **weights** and **biases** to minimize the error between predictions and actual results.
  - It is just an optimization process that can be done by using **backpropagation** and **gradient descent**.

Basic notions

The Convolution operation on an image.



check: https://hannibunny.github.io/mlbook/neuralnetworks/convolutionDemos.html

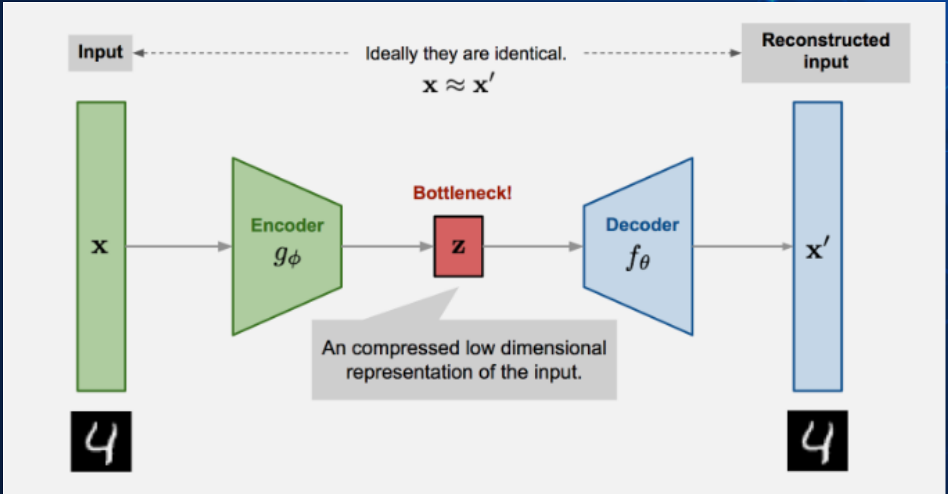
What is Unsupervised Learning?

- A type of machine learning where the model learns from data that is **not labeled**.
- **No Labels**: The data has no predefined categories or labels (like "cat" or "dog"). The algorithm has to figure out patterns on its own.
- **Discovering Patterns**: The goal is for the model to **find hidden structures** or patterns in the data. It might group similar data points together without knowing in advance what those groups represent.
- **Common Tasks**:
  - **Clustering**: Grouping data into clusters based on similarities (e.g., finding different groups of customers in a dataset).
  - **Dimensionality Reduction**: Reducing the number of features or variables while keeping the important information (e.g., simplifying image data).
- **Example**: We have multiple images from cervical MRI scans, the computer will try to group similar images together based on some discovered patterns.

Image Autoencoder for unsupervised learning

- A type of neural network used to **compress** data into a lower-dimensional form (encoding) and then **reconstruct** it (decoding).
- **Two Main Parts**:
  - **Encoder**: Compresses the input image into a smaller, lower-dimensional representation (latent space).
  - **Decoder**: Reconstructs the original image from the compressed representation.
- **Loss Function**: Measures the difference between the input image and the reconstructed image (e.g., mean squared error). The goal is to minimize this difference during training.
- **Key Idea**: Learn an efficient way to represent the image in a compressed form while being able to reconstruct the original image.
- **Why Use an Autoencoder?**
  - **Dimensionality Reduction**: Compress image data while preserving key information.
  - **Denoising**: Remove noise from images by reconstructing clean versions.
  - **Feature Learning**: The compressed representation can be useful for classification or clustering tasks.

Image Autoencoder



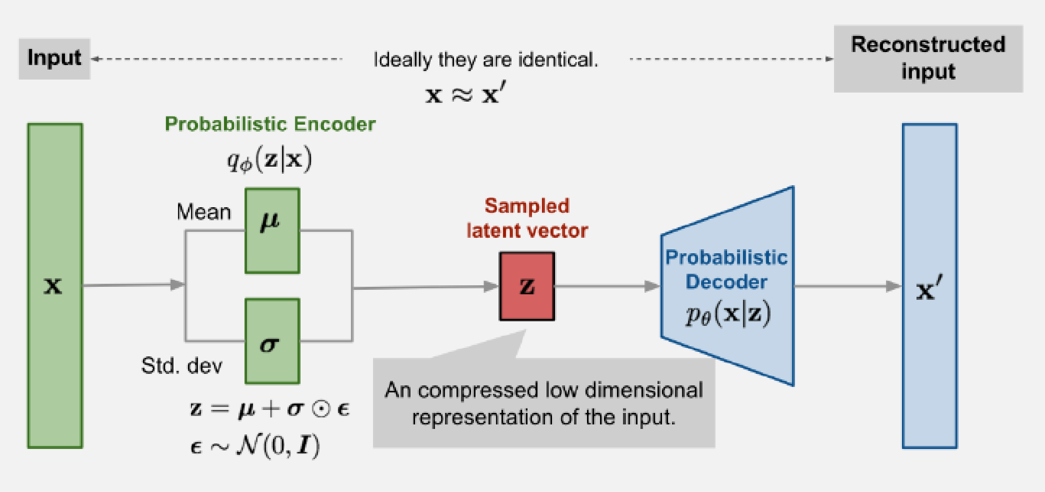
check website: https://lilianweng.github.io/posts/2018-08-12-vae/

Autoencoder vs Variational Autoencoder (VAE)

Torus AI

- **Latent Space**:
  - **Autoencoder**: Compresses input into a fixed point in latent space.
  - **VAE**: Maps the input to a probabilistic distribution in the latent space.
- **Generative Ability**:
  - **Autoencoder**: Good at reconstructing input but poor at generating new data.
  - **VAE**: Designed to generate new, realistic data by sampling from the latent space.
- **Regularization**:
  - **Autoencoder**: No explicit regularization.
  - **VAE**: Uses regularization (Kullback-Leibler divergence) to enforce structure in the latent space.
- **Loss Function**:
  - **Autoencoder**: Only uses reconstruction error (difference between input and output).
  - **VAE**: Combines reconstruction error with regularization to ensure a smooth latent space.
- **Latent Space Structure**:
  - **Autoencoder**: Can result in an unstructured latent space, making interpolation difficult.
  - **VAE**: Results in a continuous and structured latent space, allowing smooth interpolation.

Variational Autoencoder (VAE)



check website: https://lilianweng.github.io/posts/2018-08-12-vae/

Variational Autoencoder (VAE)

A VAE is a probabilistic extension of an autoencoder that enables better generative capabilities and structured latent spaces, making it useful for creating new data samples.

check website: https://taylordenouden.com/VAE-Latent-Space-Explorer/

Detect anatomic regions. Image segmentation

Torus AI

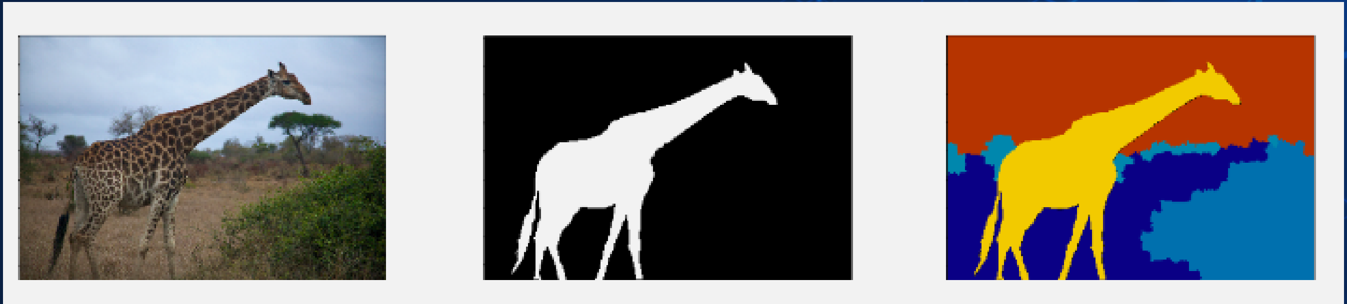
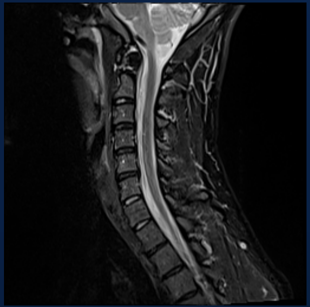


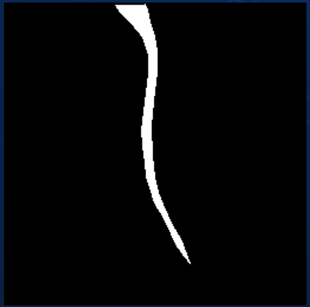
Image Segmentation

- **Image Segmentation**: A computer vision technique that divides an image into multiple regions or segments to simplify analysis.
- **Goal**:
  - To identify and separate meaningful parts (regions) of an image, such as objects, boundaries, or areas of interest.
- **How It Works**:
  - Each pixel is assigned to a segment based on criteria like color, texture, or intensity.
  - Machine learning models, especially **Convolutional Neural Networks (CNNs)**, are commonly used to perform segmentation.
- **Challenges**:
  - Requires a large amount of labeled data for supervised learning models.

Creation of the dataset

Annotated images





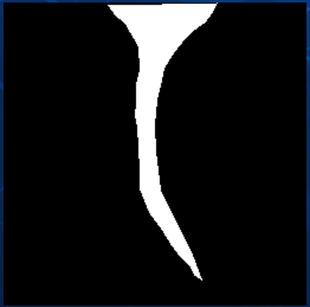
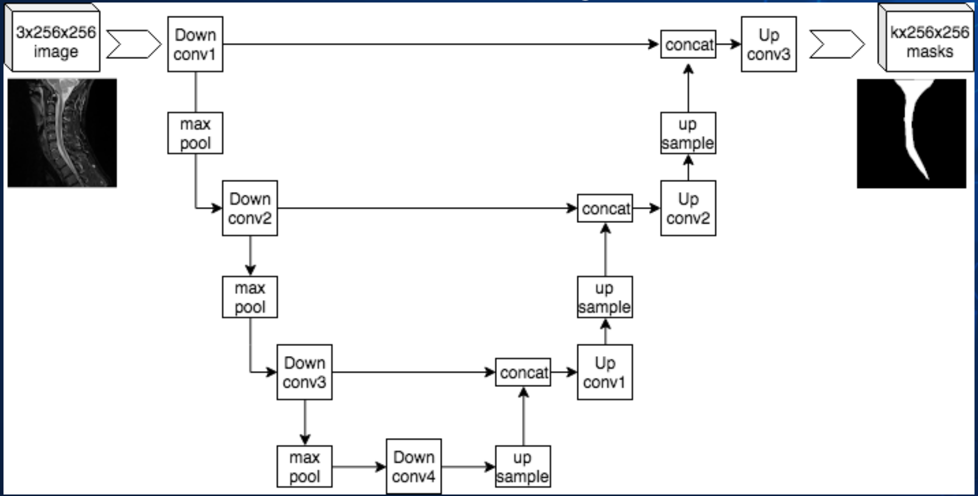


Image segmentation

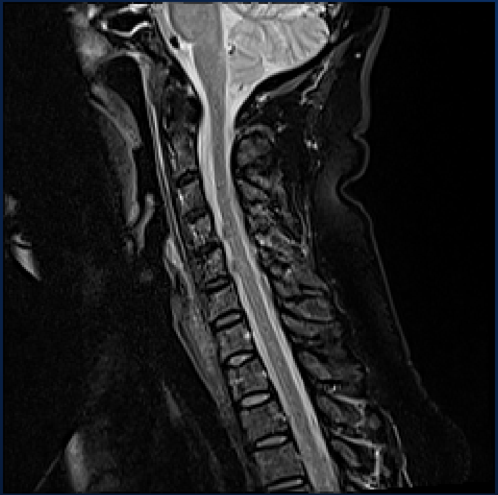
U-Net architecture for segmentation

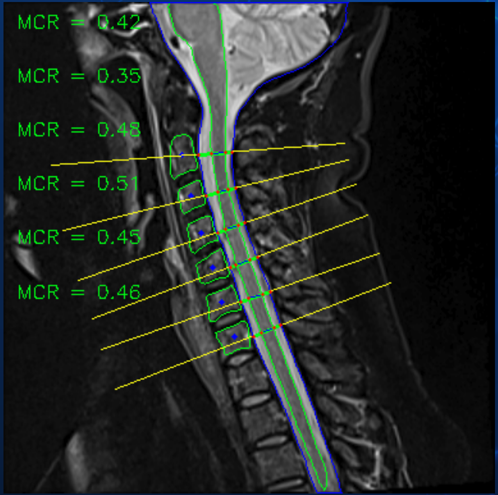


What is U-Net Architecture?

- **U-Net** is a popular neural network architecture used for **image segmentation**.
- The architecture has a **U-shaped** structure, with two main parts: the contracting (downsampling) path and the expanding (upsampling) path.
- **Contracting Path (Encoder)**: extracts features from the input image.
- **Expanding Path (Decoder)**: reconstructs the image from the compressed features.
- **Skip Connections**:
  - U-Net has **skip connections** that directly connect corresponding layers in the contracting path to layers in the expanding path.
  - This helps to preserve important spatial information that can be lost during downsampling.
- **Strengths**:
  - It works well with small datasets and produces high-quality segmentations.
  - U-Net captures both the context (global) and the fine details (local) of the image.

The Complete Solution





Tool used for the solution

- **Deep Learning architectures**
  - Variational Autoencoder (VAE) for unsupervised learning
  - U-Net architecture for image segmentation

- **Python libraries**:
  - pydicom (medical images, specially DICOM format)
  - pytorch (deep learning architectures and optimization)
  - numpy (post processing)
  - opencv-python (post processing)
  - PIL (post processing)

Future Directions and Challenges

Torus AI

Retrospective clinical studies involving other medical doctors

Certification and Validation (e.g. FDA validation)

Integration into professional medical software

Thanks !