FAKULTAS ILMU KOMPUTER

Semester Gasal 2024/2025
Topic #13

Recent Topics in Computer Vision: Biomedical Image Analysis using CV & DL

CSCE604133 Computer Vision
Faculty of Computer Science
Universitas Indonesia

Muhammad Febrian Rachmadi, Ph.D.
Dr. Eng. Laksmita Rahadianti
Dr. Dina Chahyati, Prof. Dr. Aniati M. Arymurthy

Recall: The 3R of Computer Vision

- **Recognition:** Highlighting objects-of-interest of a specific task.
  - Examples: Classification, detection, segmentation
- **Reconstruction:** Improving information contained inside an image.
  - Examples: Denoising, super-resolution, 3D reconstruction

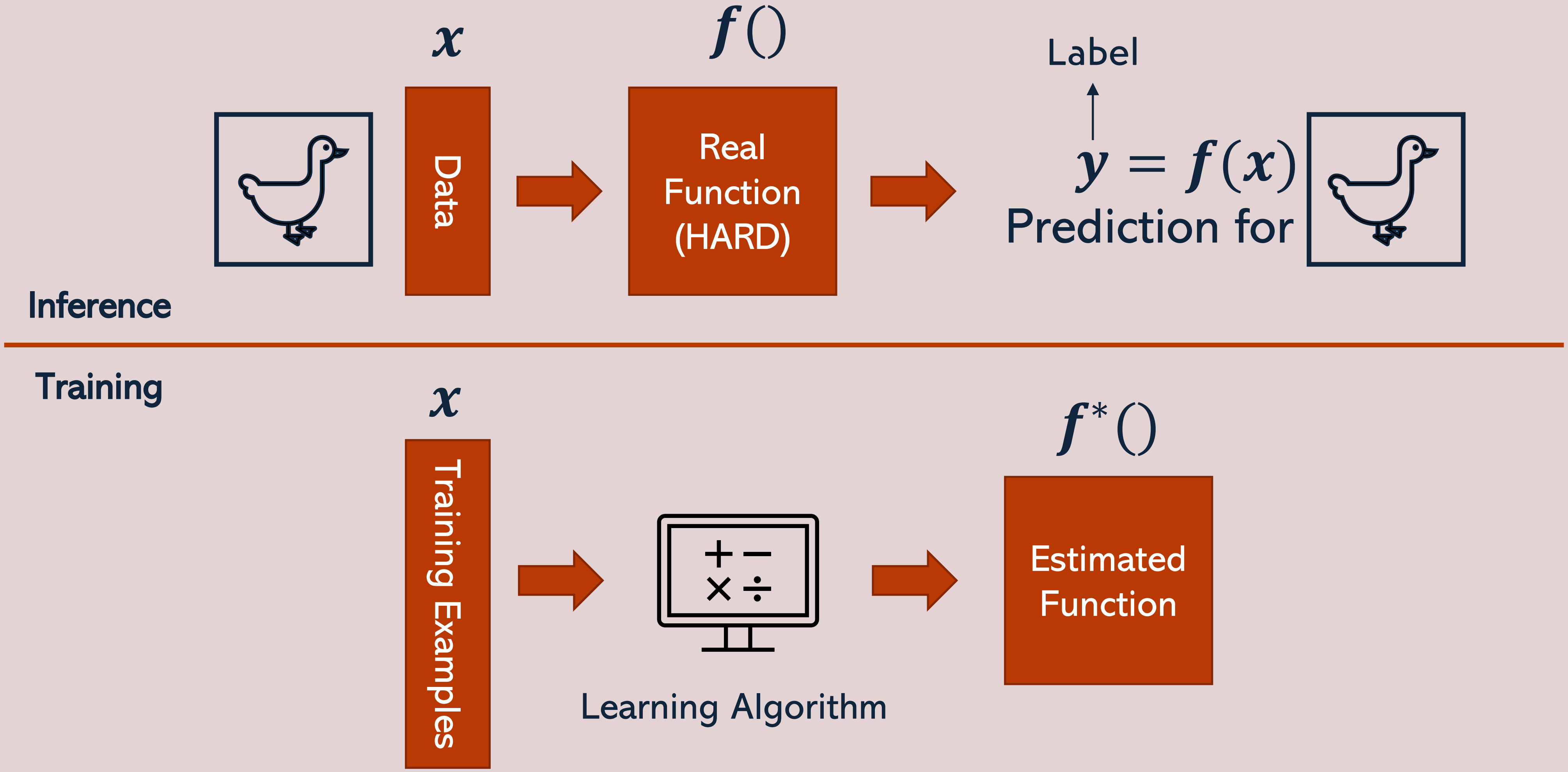
biar bisa analisa lebih baik

- **Reorganization:** Spatially move an image to another image's space
  - Examples: Registration, stitching, merging, fusing

mengangkat resolusi dari sebuah image

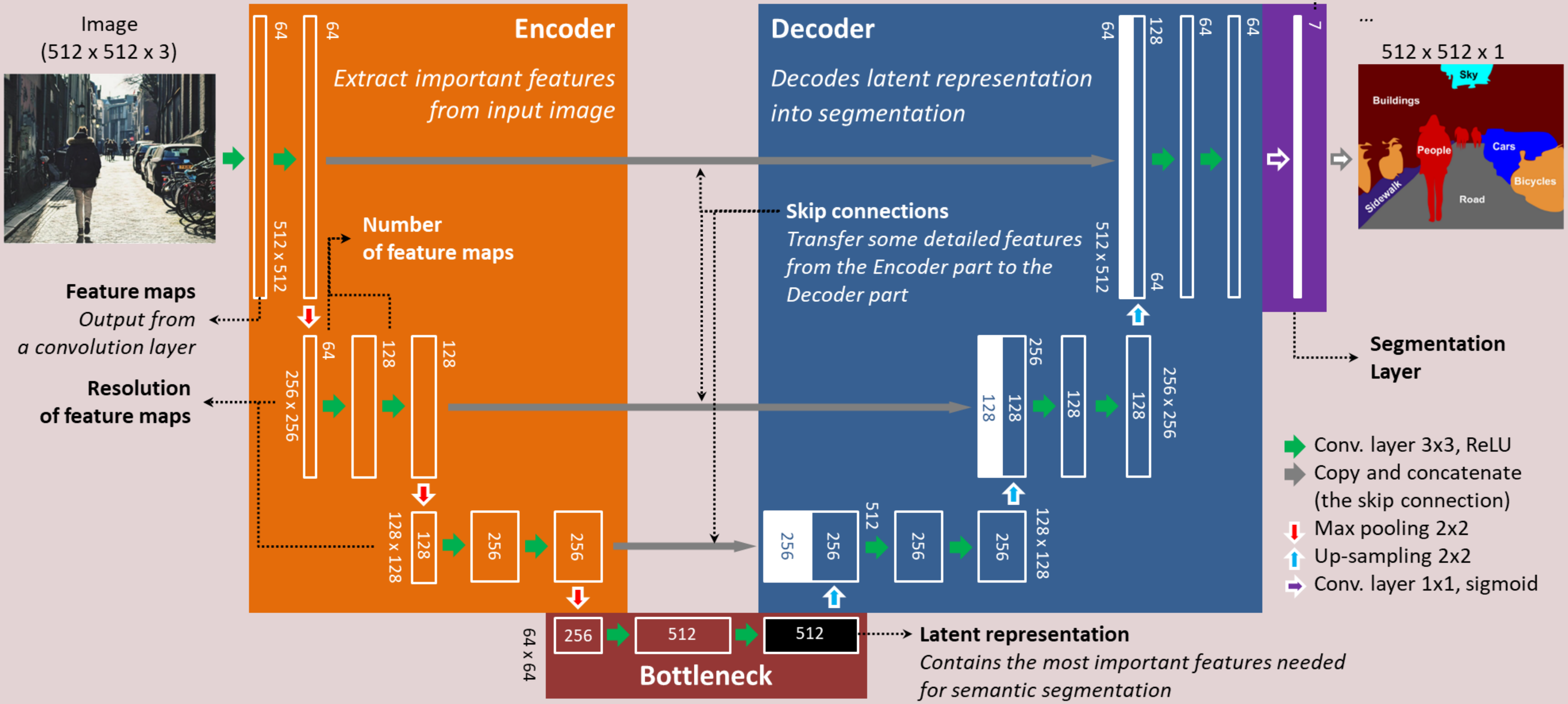
- **Why use ML/DL for biomedical image analysis?**
  - Compared with natural images, biomedical images are harder to analyze (especially 3D images, such as CT/MRI, and videos, such as ultrasound/USG).
  - Computationally expensive due to ill-posed problems (especially in reconstruction and reorganization tasks).
- So, all tasks of the 3R of CV are easier when ML/DL methods are used.

Learning from Examples using ML/DL



The Importance of Biomedical Image Analysis

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. **"U-net: Convolutional networks for biomedical image segmentation."** *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.



FAKULTAS ILMU KOMPUTER
UNIVERSITAS INDONESIA

Agenda

- Recognition in Biomedical Image Analysis
- Reconstruction in Biomedical Image Analysis
- Reorganization in Biomedical Image Analysis
- Explainable Artificial Intelligence (XAI)
- Uncertainty in Biomedical Image Analysis

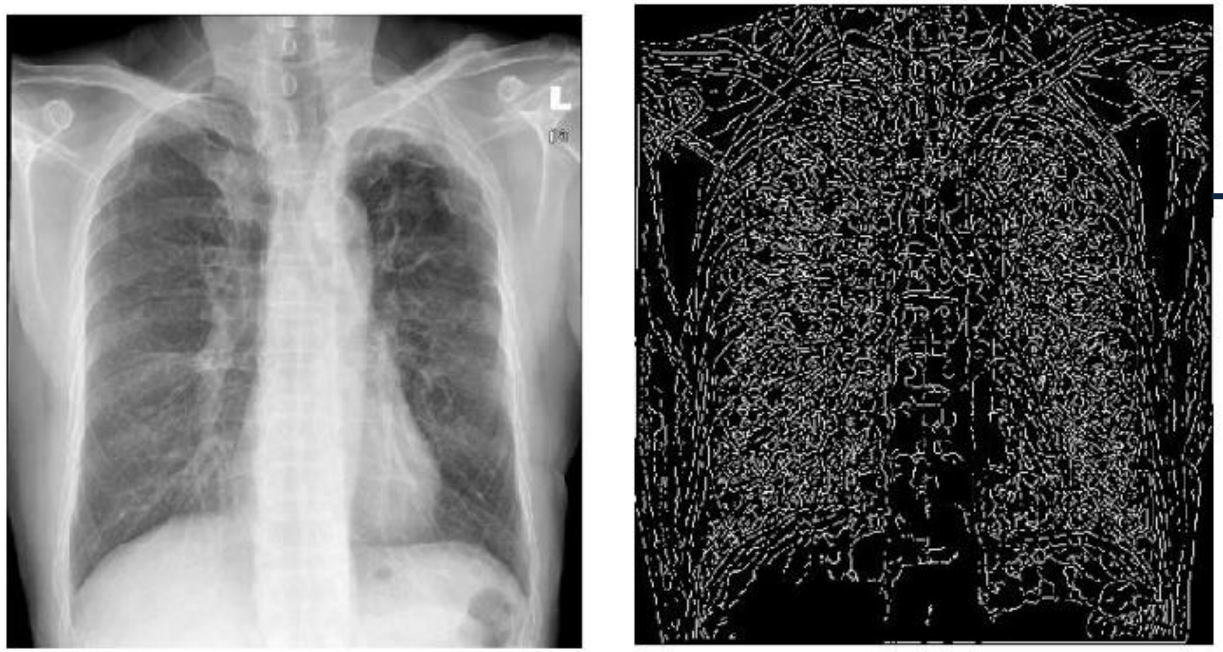
Recognition in Biomedical Image Analysis

Section 1

Case study: Chest X-ray Classification Using Deep Learning for Automated COVID-19 Screening

Chest X-ray without Deep Learning

- It is hard to extract features from low-resolution and noisy images such as chest X-rays without deep learning.



(A) (B)

Fig. 3. (A) the original image and (B) the edge detected

membutuhkan deep learning buat ngertiin chest x ray, usually human masih kesulitan untuk klasifikasi apapun

The extracted features do not describe anything useful…

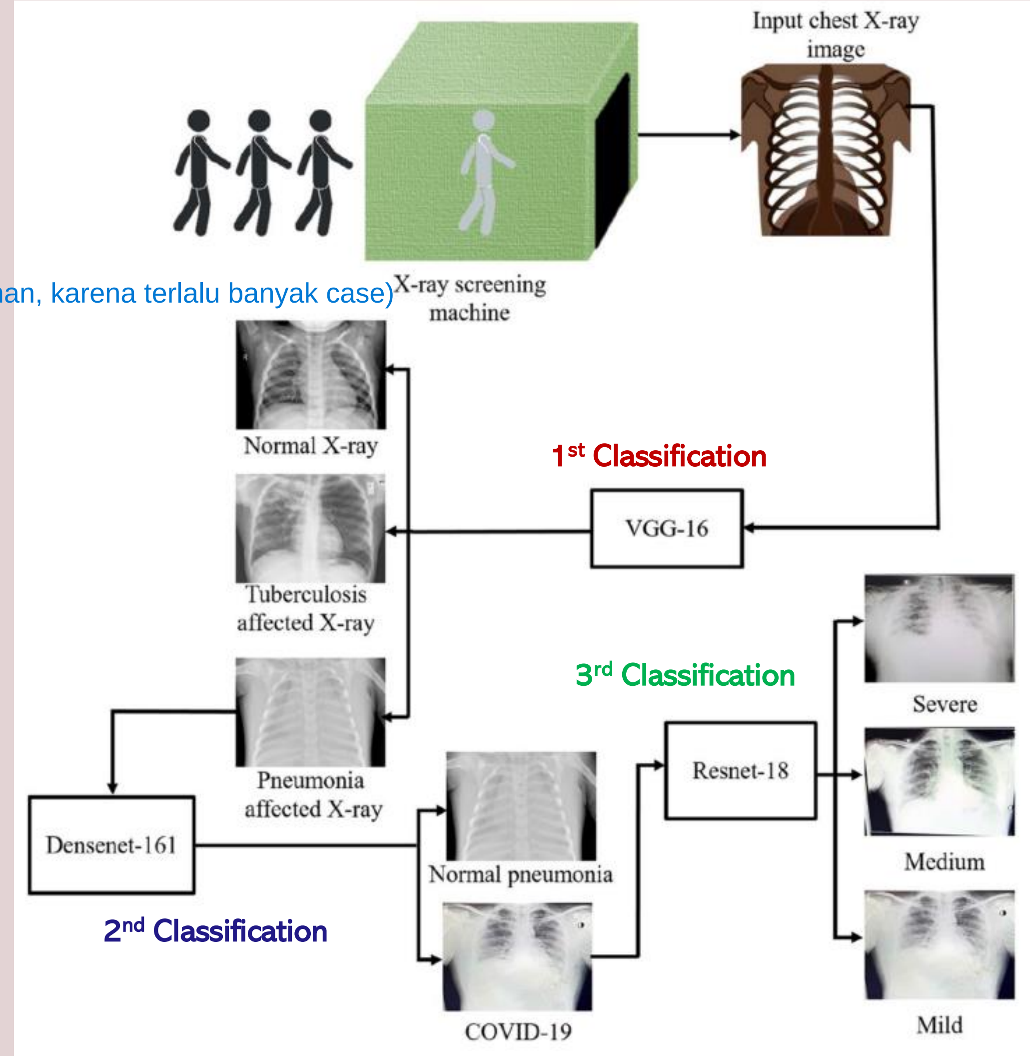
Saad, Mohd Nizam, et al. "Image segmentation for lung region in chest X-ray images using edge detection and morphology." *2014 IEEE international conference on control system, computing and engineering (ICCSCE 2014)*. IEEE, 2014.

The Proposed Method

- A cascaded classification method using different DL classification methods for 4 classes, i.e., Normal, TB, Pneumonia, and COVID-19, labeled by 3 experts.
  - Why cascade? Simplification of the problem and reducing uncertainty in each case.

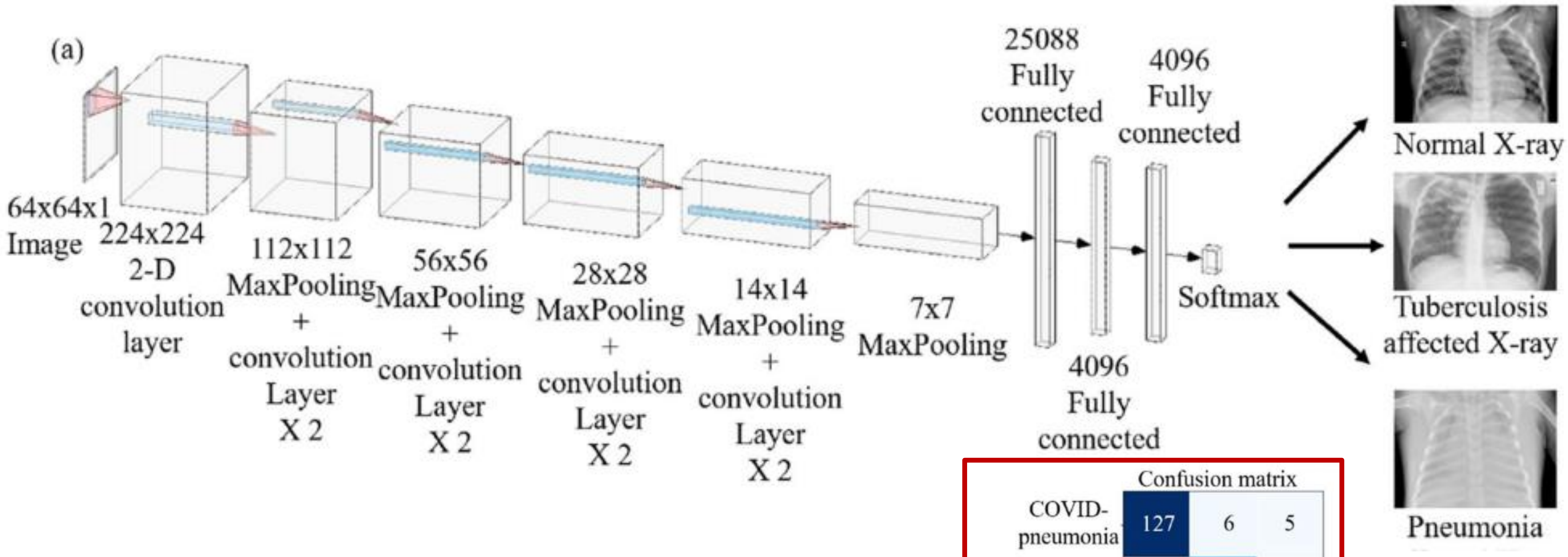
perlu dipisahin karena biar
mengurangi uncertainity (ketidakyakinan, karena terlalu banyak case)

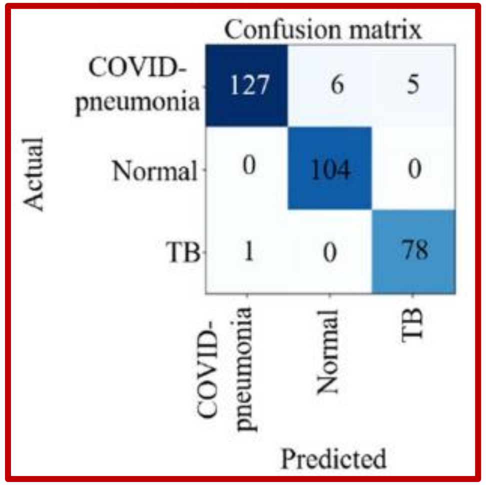
- The experts were qualified with a postgraduate degree of Doctor of Medicine (MD) in Radiology.
  - These experts independently assessed the data and in case of any disagreement, the class label received from two out of the three experts was considered.
  - No case arose where all three experts gave a different opinion.
- A total of 2271 chest X-ray images (895 × 1024 × 3 pixels) were obtained from Clinico Diagnostic Lab, Mumbai, India.



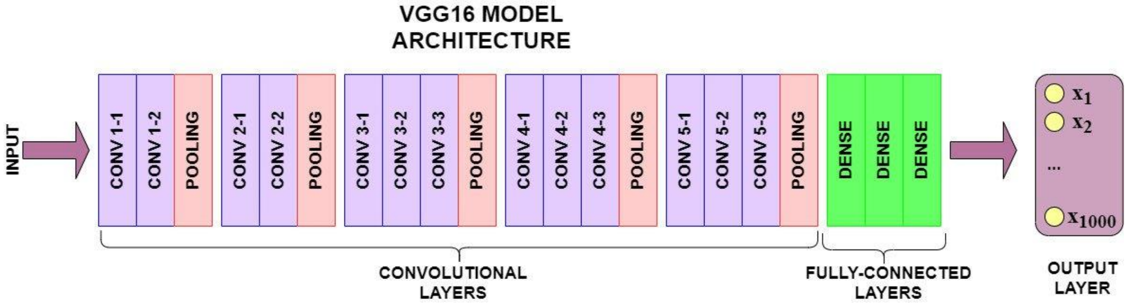
Shelke, Ankita, et al. "Chest X-ray classification using deep learning for automated COVID-19 screening." *SN computer science* 2.4 (2021): 300.

1st Classification using VGG-16



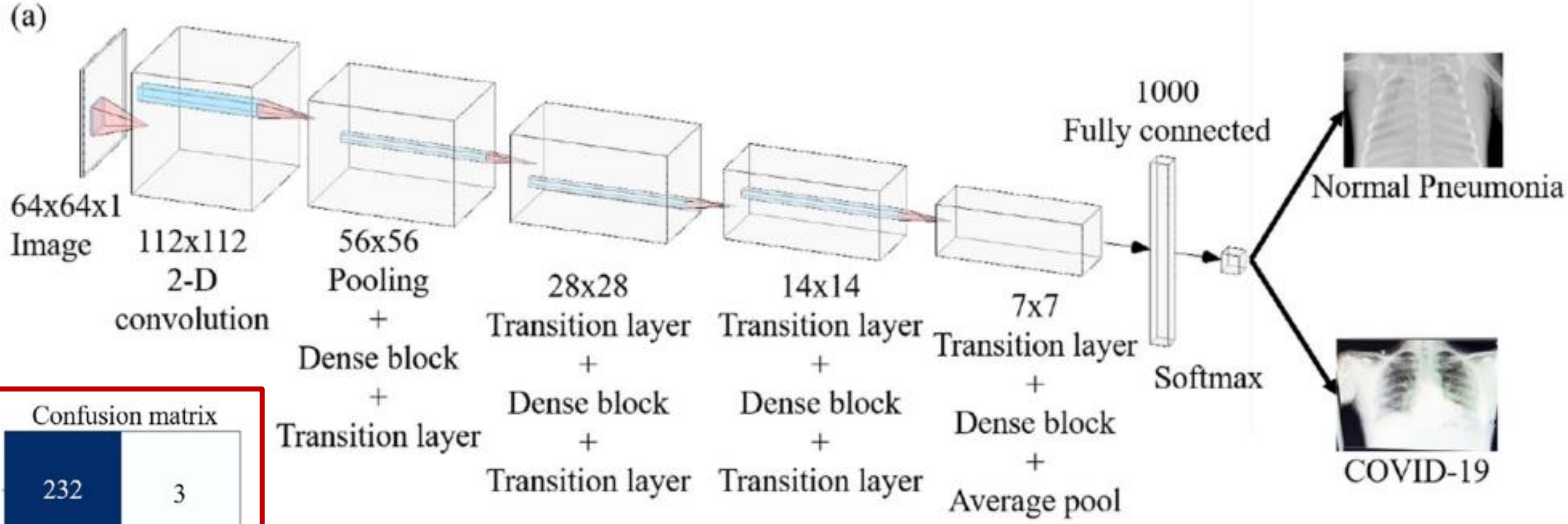


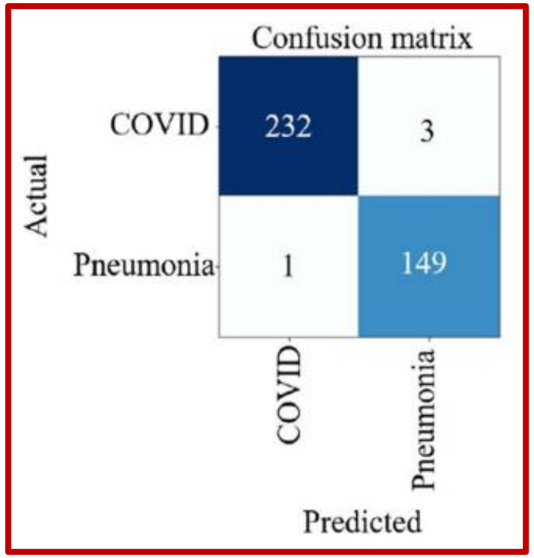
https://www.learndatasci.com/tutorials/hands-on-transfer-learning-keras/

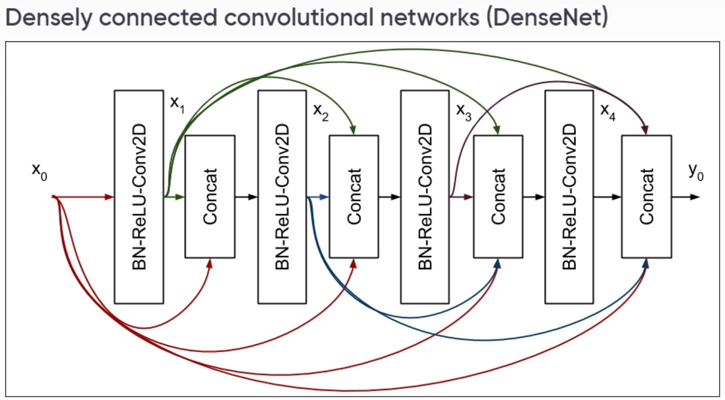


Shelke, Ankita, et al. "Chest X-ray classification using deep learning for automated COVID-19 screening." *SN computer science* 2.4 (2021): 300.

2nd Classification using DenseNet-161



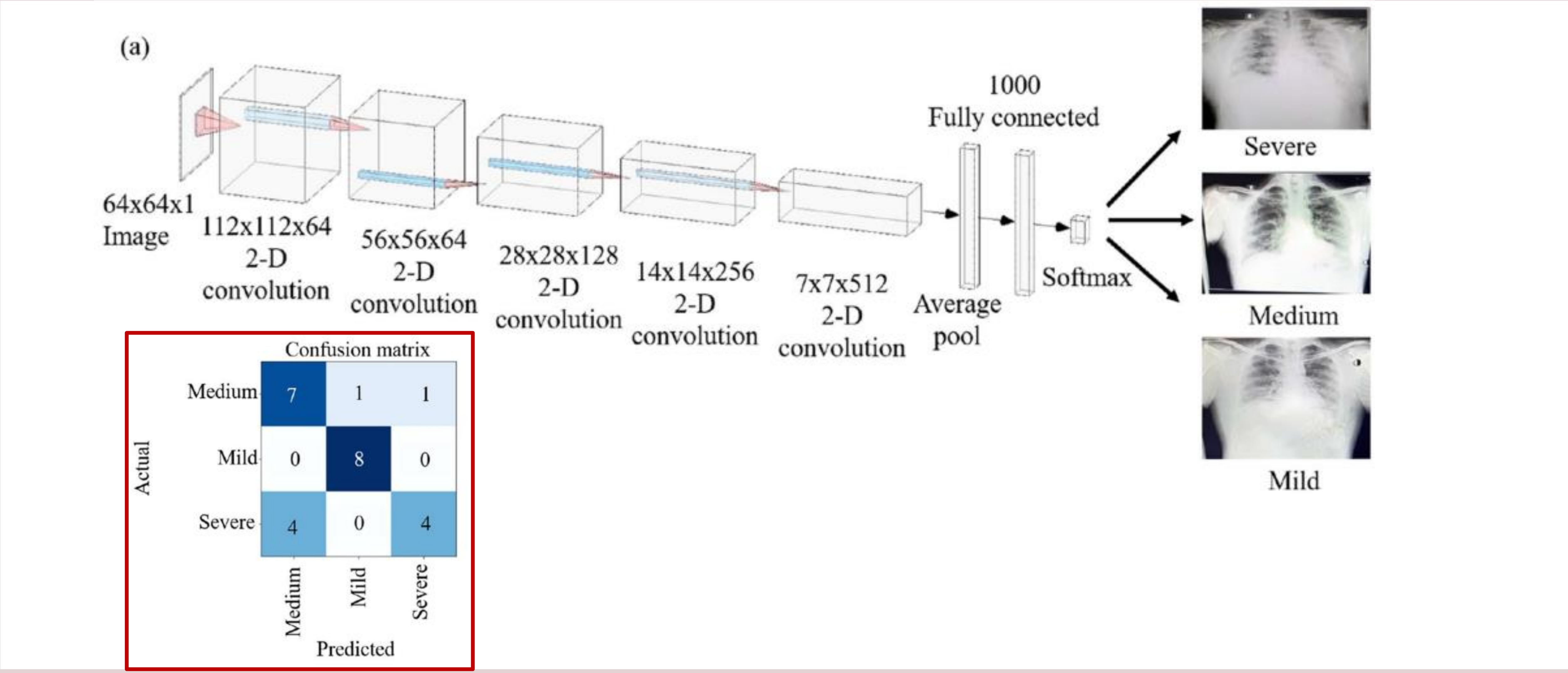




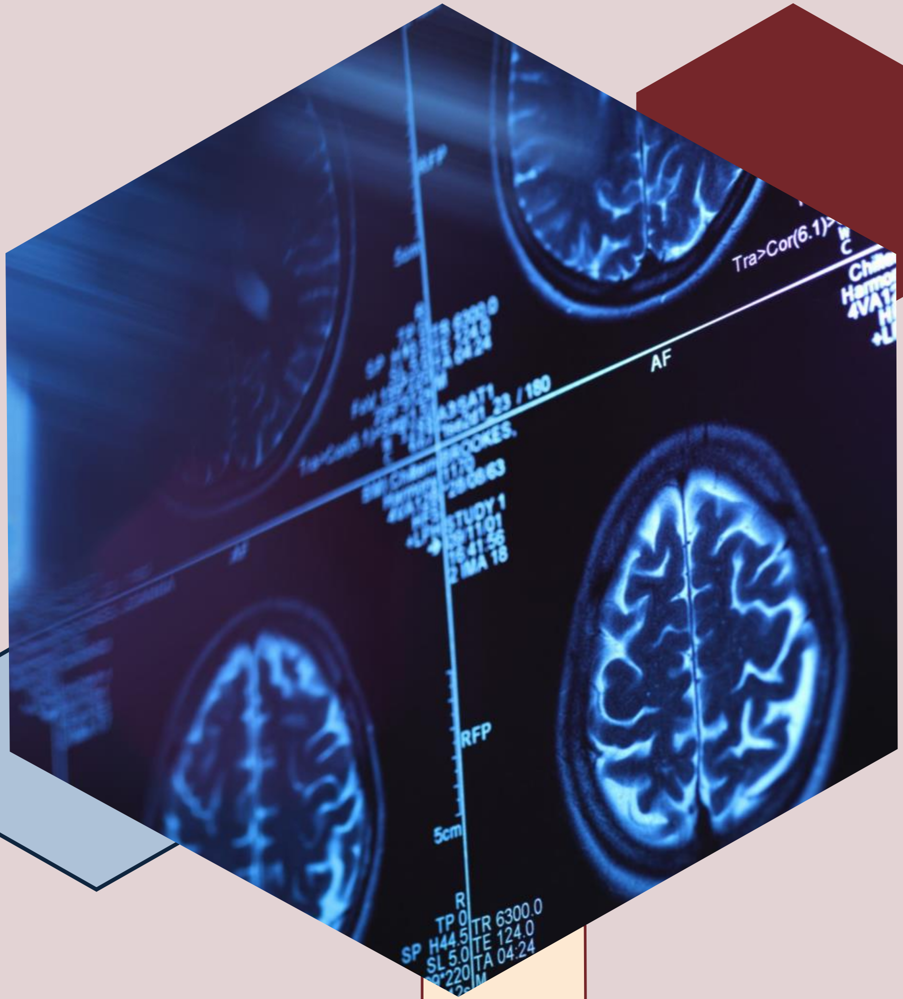
https://www.oreilly.com/library/view/advanced-deep-learning/9781788629416/ch02s04.html

Shelke, Ankita, et al. "Chest X-ray classification using deep learning for automated COVID-19 screening." *SN computer science* 2.4 (2021): 300.

3rd Classification using ResNet-16



Shelke, Ankita, et al. "Chest X-ray classification using deep learning for automated COVID-19 screening." *SN computer science* 2.4 (2021): 300.



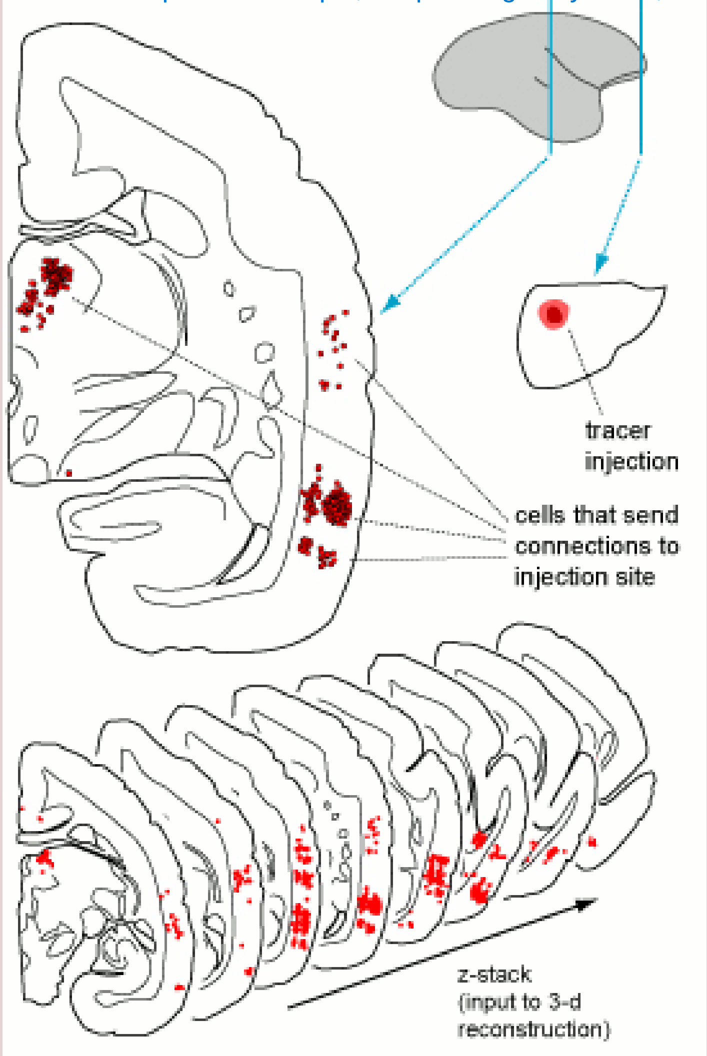
Reconstruction in Biomedical Image Analysis

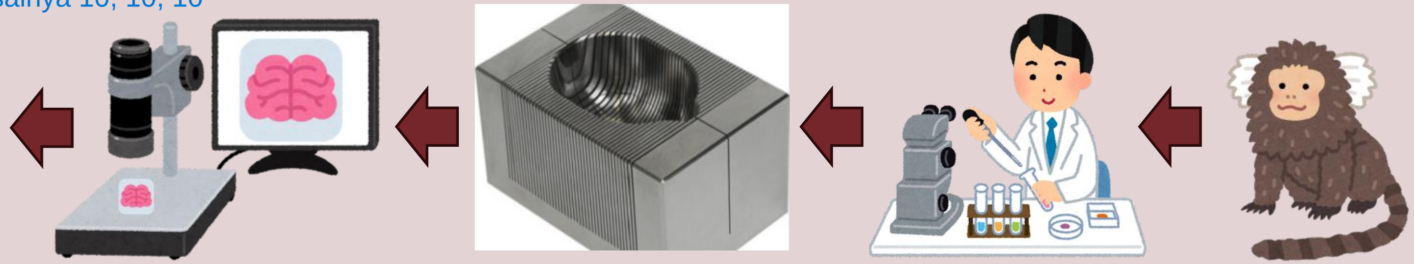
Section 2

Case study: DiffuseIR: Diffusion Models For Isotropic Reconstruction of 3D Microscopic Images

Anisotropic Spatial Resolution of Microscope Image

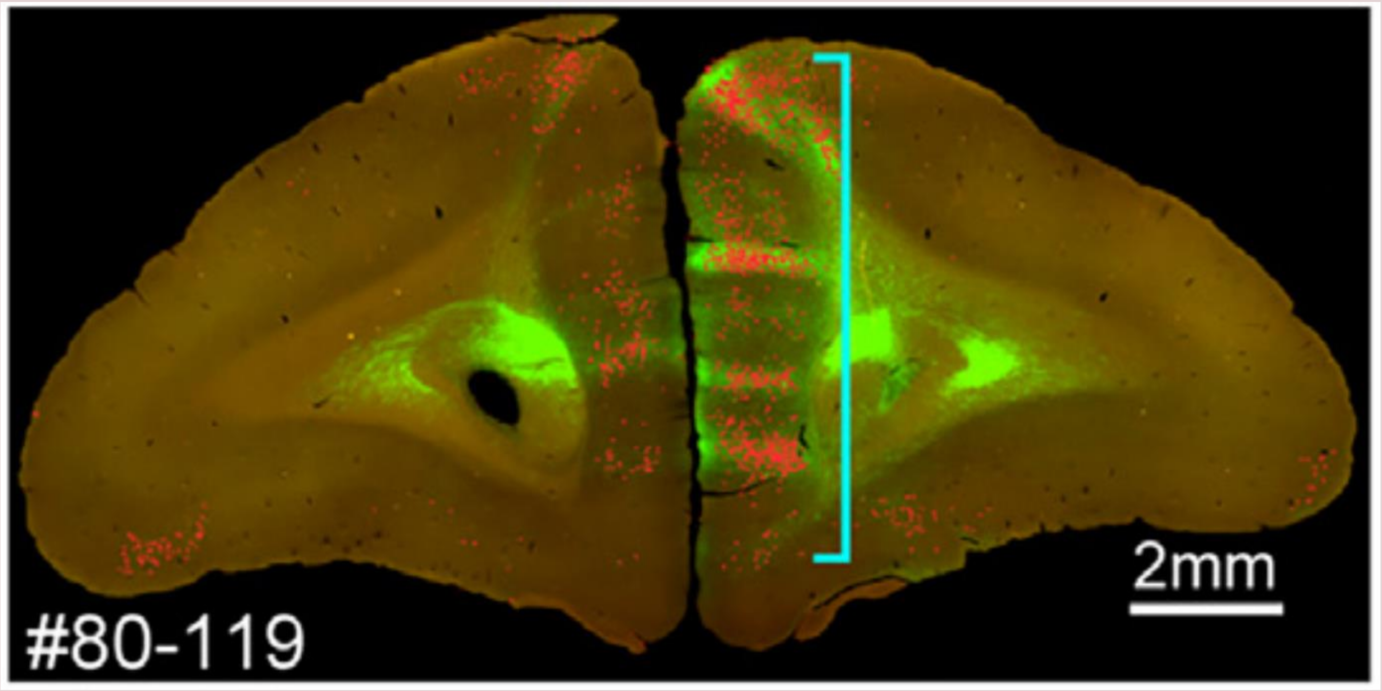
dia itu pembagiannya gak rata, misalnya 10, 10, 20
tapi kalo isotropic, dia pembagiannya rata, misalnya 10, 10, 10





High Resolution in the x-axis (thousands of pixels)

High Resolution in y-axis (thousands of pixels)



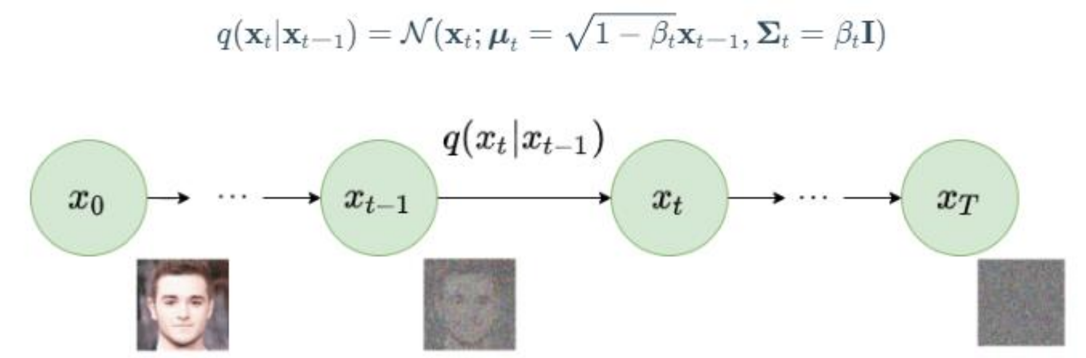
The z-stack has a low resolution (only hundreds of sections), which creates an **anisotropic 3D microscopic image**.

Watakabe, Akiya, Henrik Skibbe, Ken Nakae, Hiroshi Abe, Noritaka Ichinohe, Muhammad Febrian Rachmadi, Jian Wang et al. "Local and long-distance organization of prefrontal cortex circuits in the marmoset brain." Neuron 111, no. 14 (2023): 2258-2273.

DiffuseIR: Diffusion Models for Isotropic Reconstruction of 3D Microscopic Images

- Three-dimensional microscopy is often limited by anisotropic spatial resolution, resulting in lower axial resolution (i.e., the z-stack) than lateral resolution (i.e., x-axis and y-axis).
- To address these issues, we propose DiffuseIR, an unsupervised method for isotropic reconstruction based on diffusion models. **Diffusion models** are a new class of state-of-the-art generative models that generate diverse high-resolution images. Untuk membuat image yang berapa di gap (yang membutuhkan isotropic reconstruction)

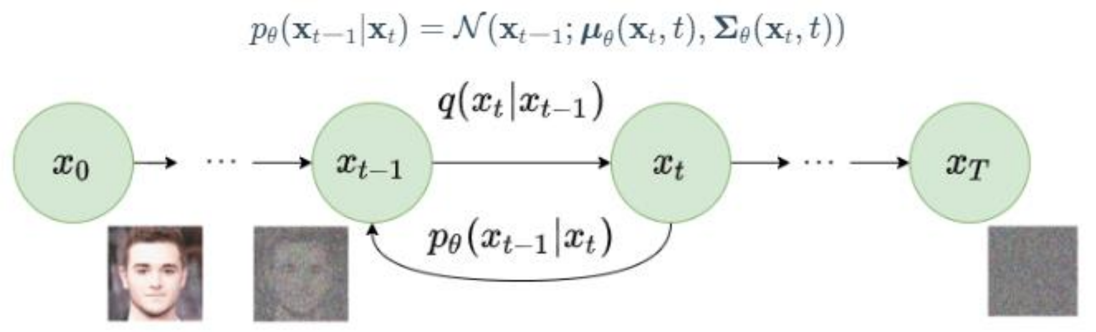
1. **Forward diffusion:** Take an input image x_0 and gradually add T steps of Gaussian noise.



$$\begin{aligned}\mathbf{x}_t &= \sqrt{1-\beta_t}\mathbf{x}_{t-1} + \sqrt{\beta_t}\boldsymbol{\epsilon}_{t-1}\\ &= \sqrt{\alpha_t}\mathbf{x}_{t-2} + \sqrt{1-\alpha_t}\boldsymbol{\epsilon}_{t-2}\\ &= \dots\\ &= \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}_0\end{aligned}$$

$$\mathbf{x}_t \sim q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I})$$

2. **Reverse diffusion:**

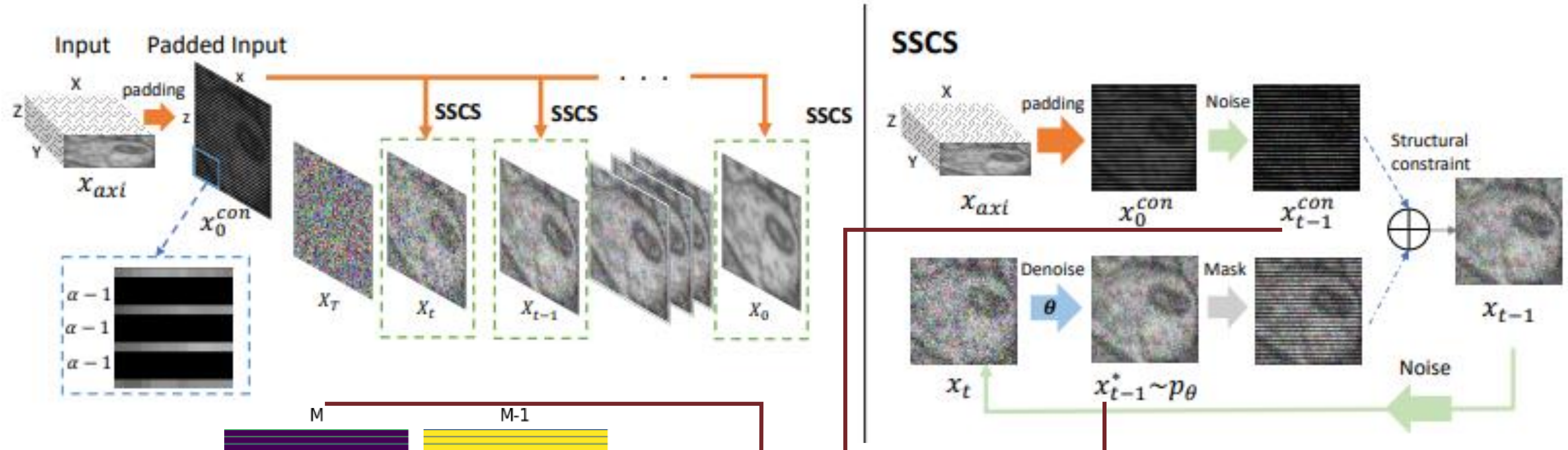


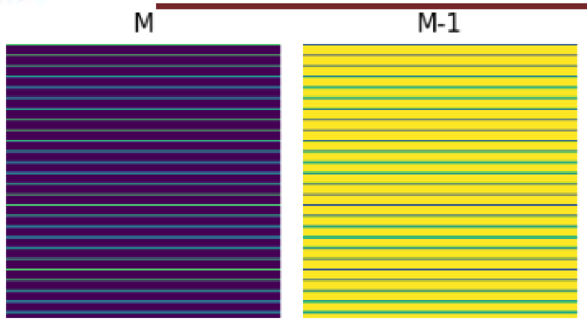
Training a diffusion model: https://theaisummer.com/diffusion-models/

Pan, Mingjie, et al. "DiffuseIR: Diffusion models for isotropic reconstruction of 3d microscopic images." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer Nature Switzerland, 2023.

DiffuseIR's Sparse Spatial Condition Sampling (SSCS)

- SSCS extracts sparse structure context from low axial-resolution slice x_{axi} and generates reconstruction result $x_0 \sim p_\theta(x_{lat}|x_{axi})$.





$$x_{t-1} = M \odot x_{t-1}^{con} + (1 - M) \odot x_{t-1}^{*}) \tag{4}$$

Reconstruction Results

CV | GAN-based | Supervised | Diffusion-based

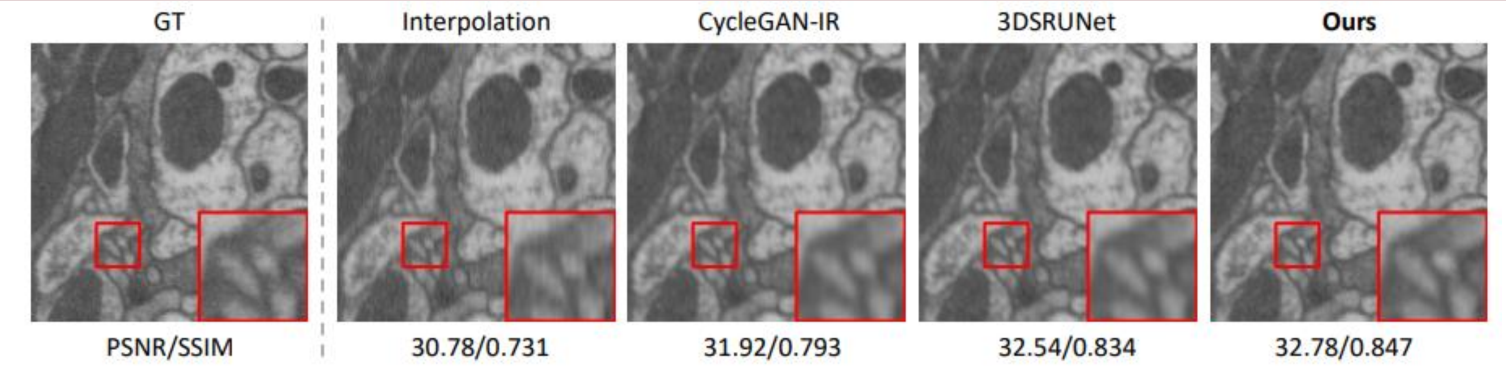
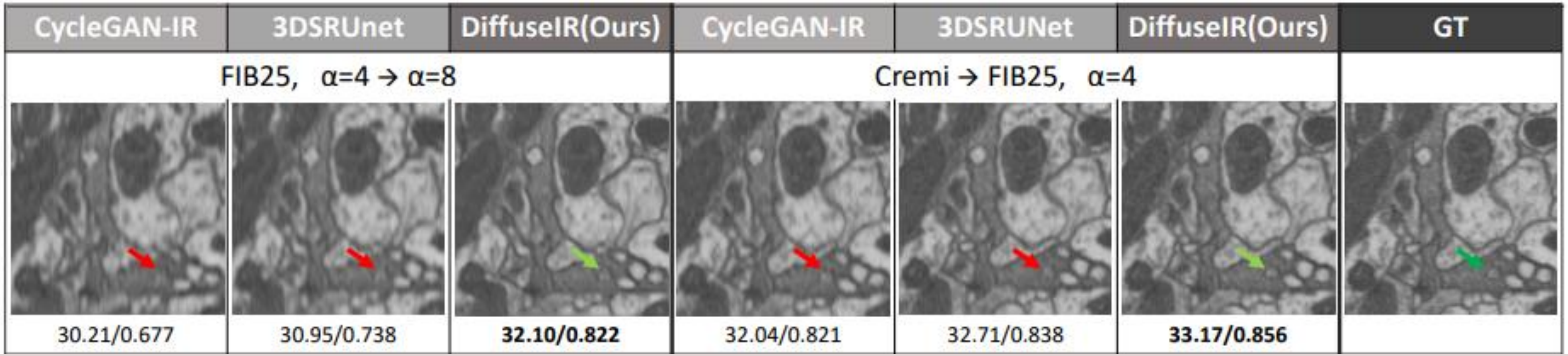


Fig. 2: **Visual comparisons on FIB-25 dataset ($\alpha = 4$).** DiffuseIR can generate competitive results compared to supervised methods, and the results appear more visually realistic.



- Some measurements in reconstruction:
  - **PSNR:** Peak signal-to-noise ratio, where higher values are better.
  - **SSIM:** Structural similarity index measure, where higher values are better.
- Reconstruction is hard to quantitatively measure, especially for measuring quality as perceived by humans.

Reorganization/Registration in Biomedical Image Analysis

Section 3

Case study: Voxelmorph: A learning framework for deformable medical image registration

Image Registration

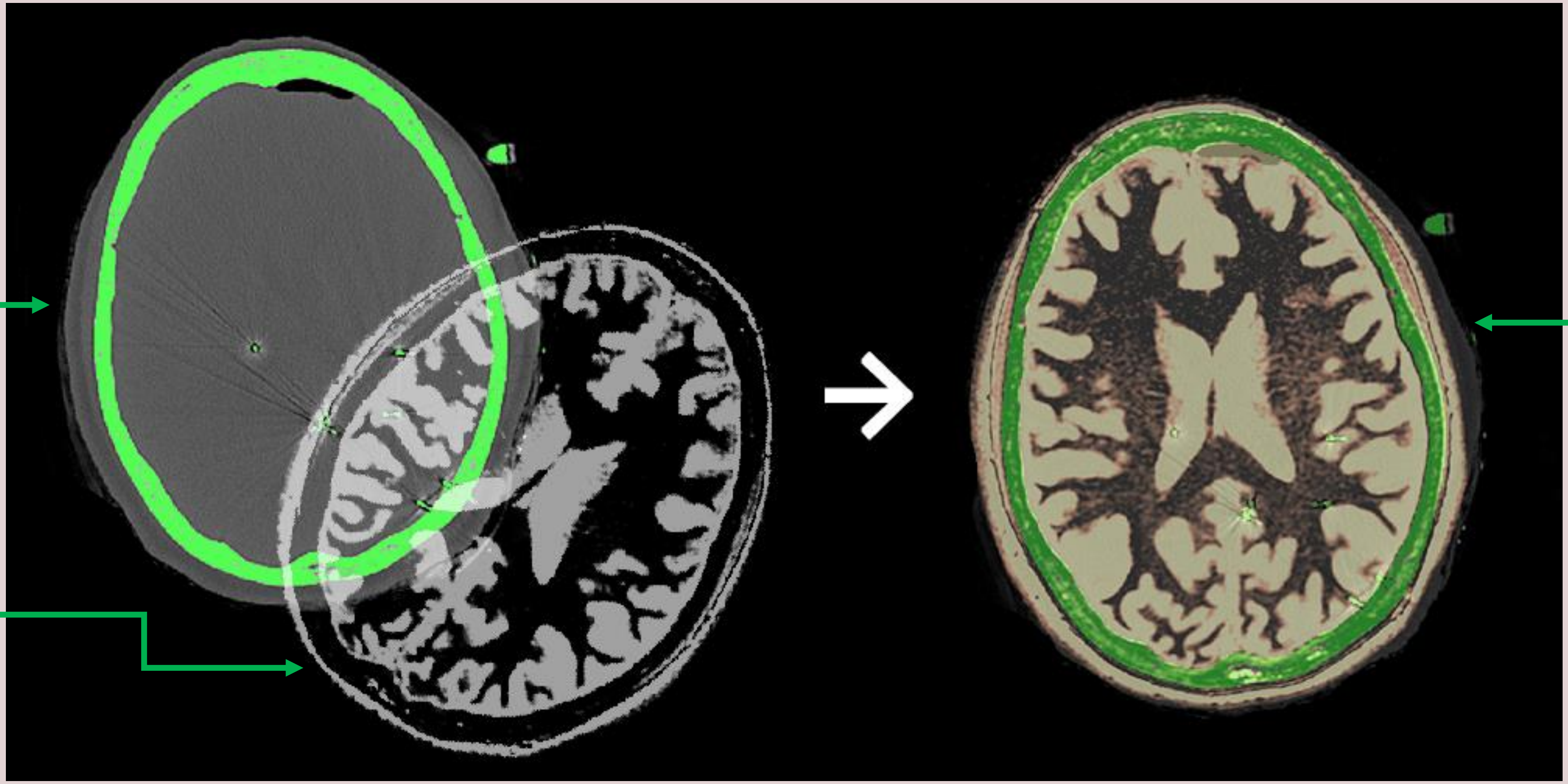
membandingkan visual dari CT Scan dengan MRI

tapi sekarang kedua scan itu, hasilnya berbeda. Nah jadi kita membutuhkan registration, dimana MRI align dengan CT Scan

- **Registration** is the process of establishing spatial alignment between two images in two different spaces. It allows for the alignment and transfer of key information across subjects and atlases/references.
- One of the hardest and most challenging tasks in biomedical image analysis due to an ill-posed problem!

CT scan is used as the reference image/space

MRI is used as the moving image



The CT and MR imaging data (left) are unaligned before registration. The aligned data (right) show the correct anatomical positioning after registration. (https://3dqlab.stanford.edu/image-registration/)

Several transformations are usually needed in the registration process:

1. Rigid transformation
2. Affine transformation
3. Deformable transformation

These transformations could be very complex for high-resolution and 3D images. That is when the DL-based registration method shines compared to the statistical-based registration method.

Chen, Min, et al. "Image registration: fundamentals and recent advances based on deep learning." *Machine Learning for Brain Disorders* (2023): 435-458.

VoxelMorph: The Baseline for Image Registration

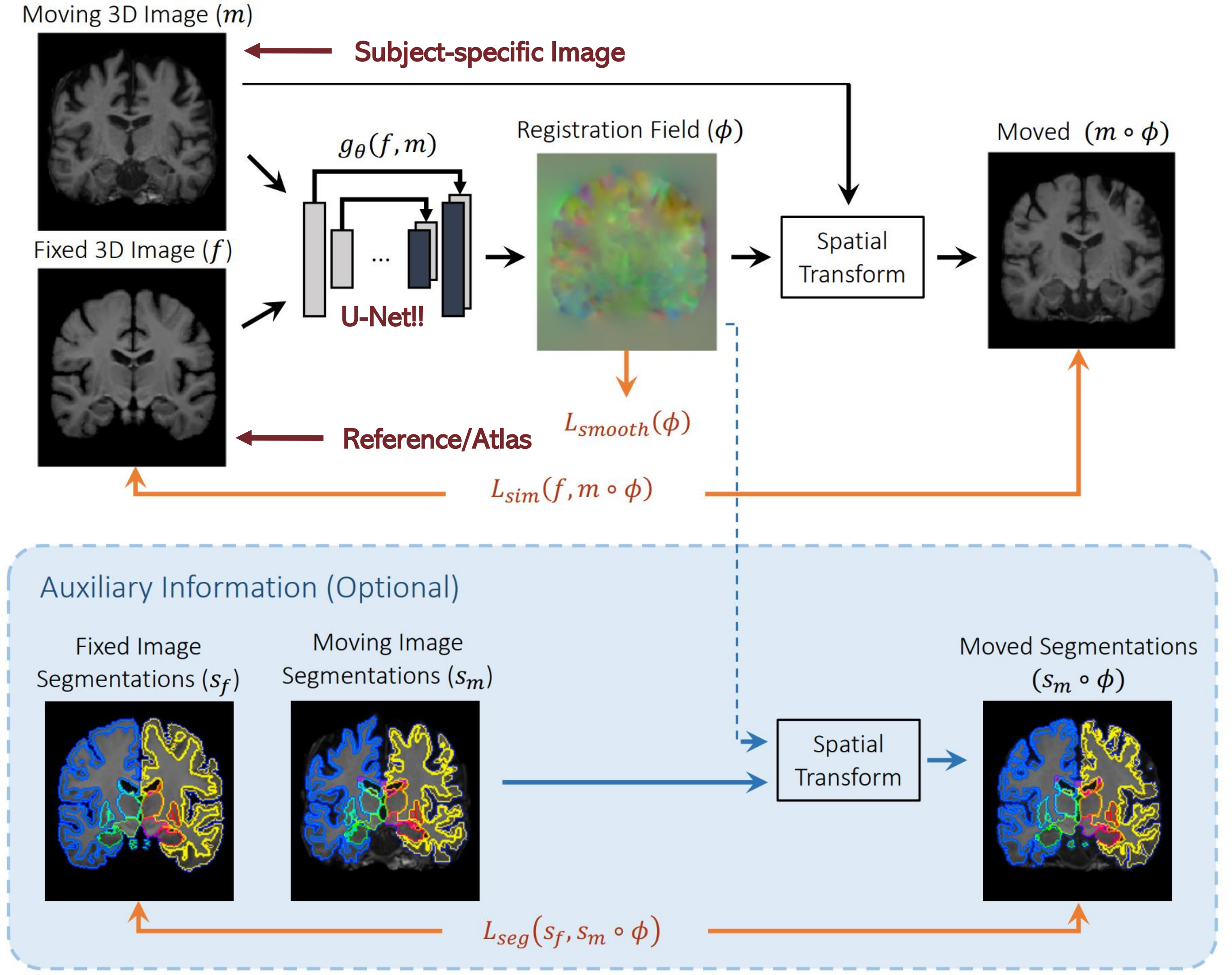
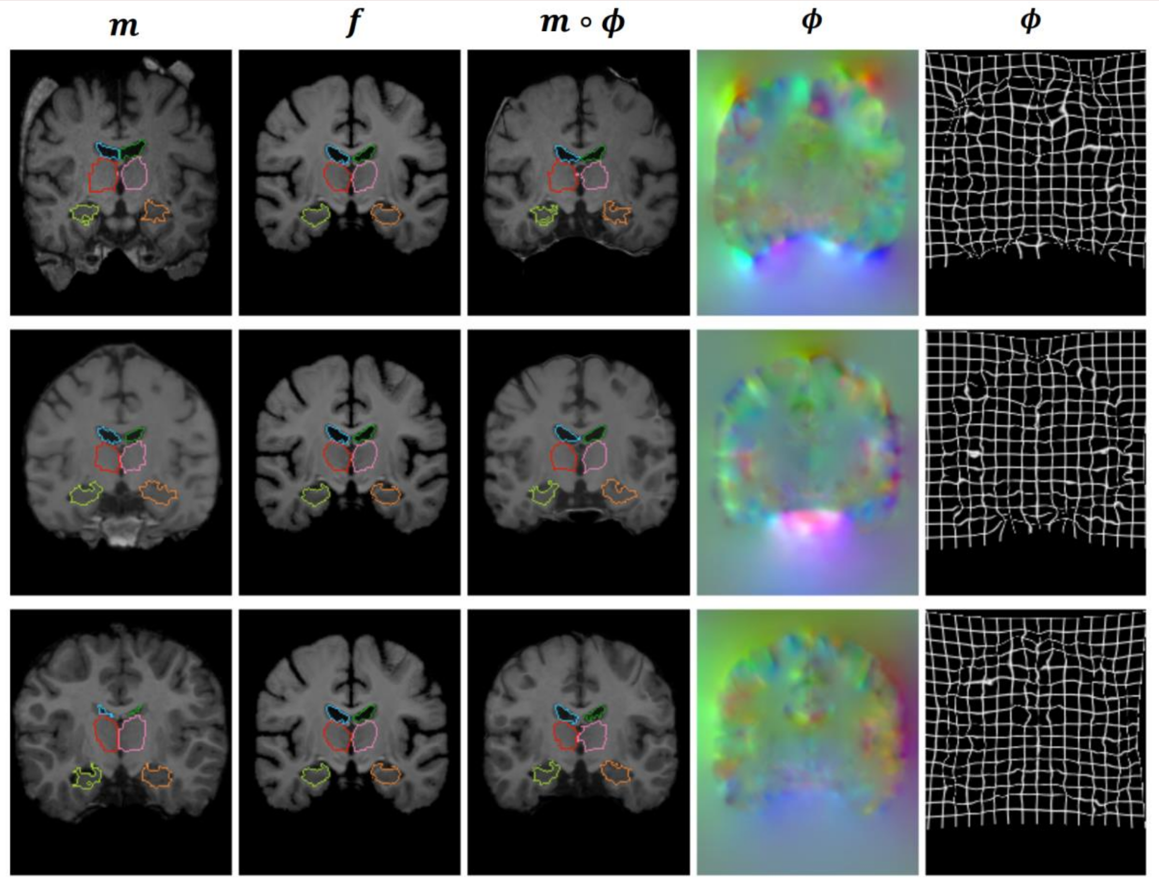
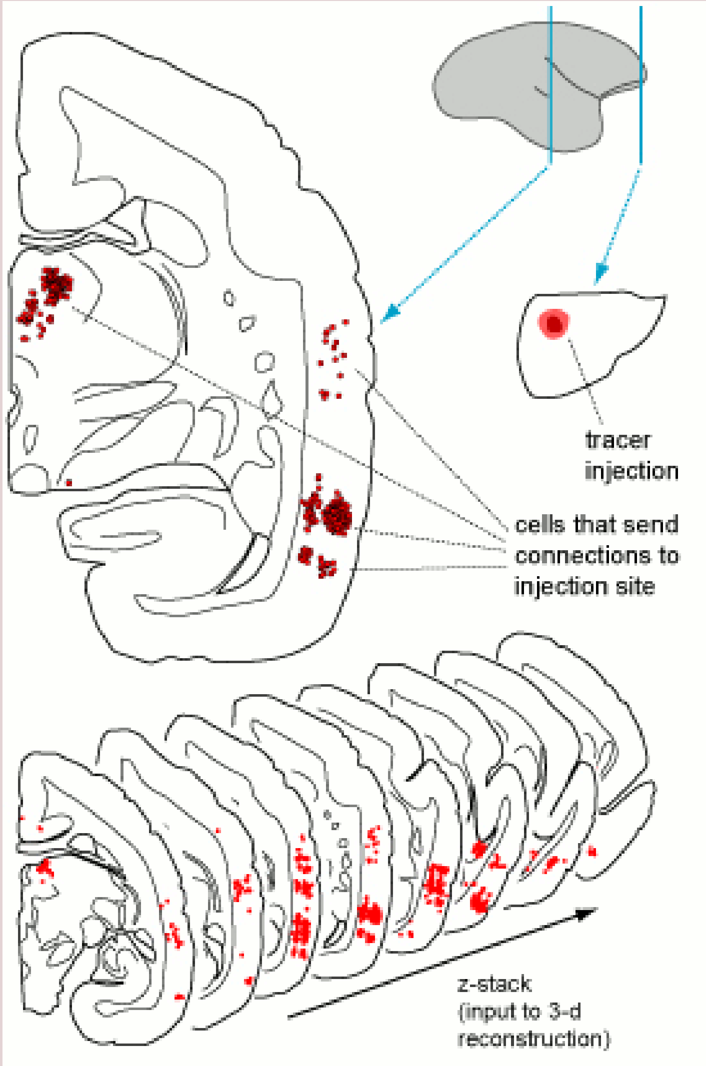


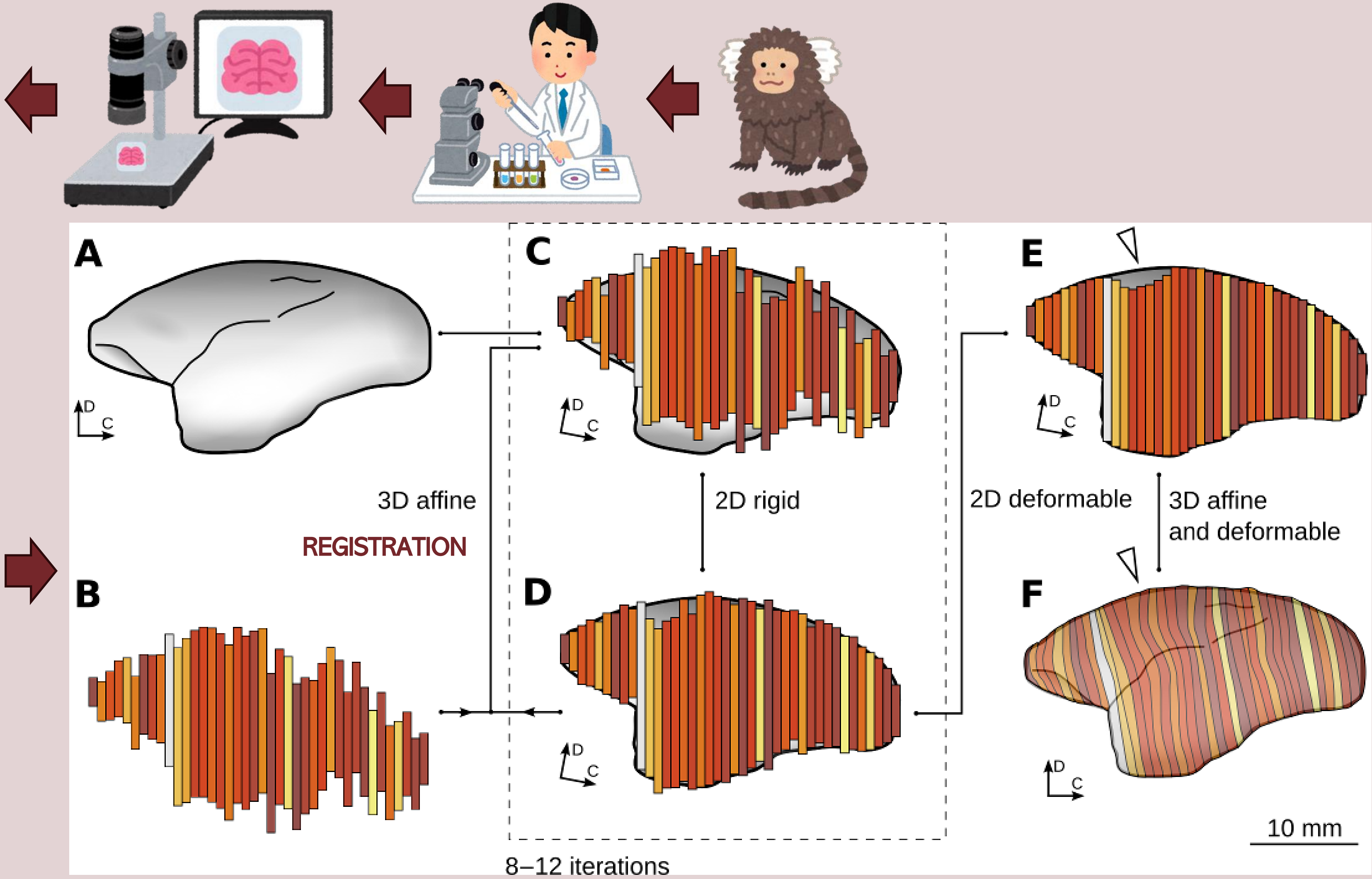
Fig. 2: Overview of the method. We learn parameters θ for a function $g_\theta(\cdot,\cdot)$, and register 3D volume m to a second, fixed volume f. During training, we warp m with ϕ using a spatial transformer function. Optionally, auxiliary information such as anatomical segmentations $\mathbf{s_f}, \mathbf{s_m}$ can be leveraged during training (blue box).



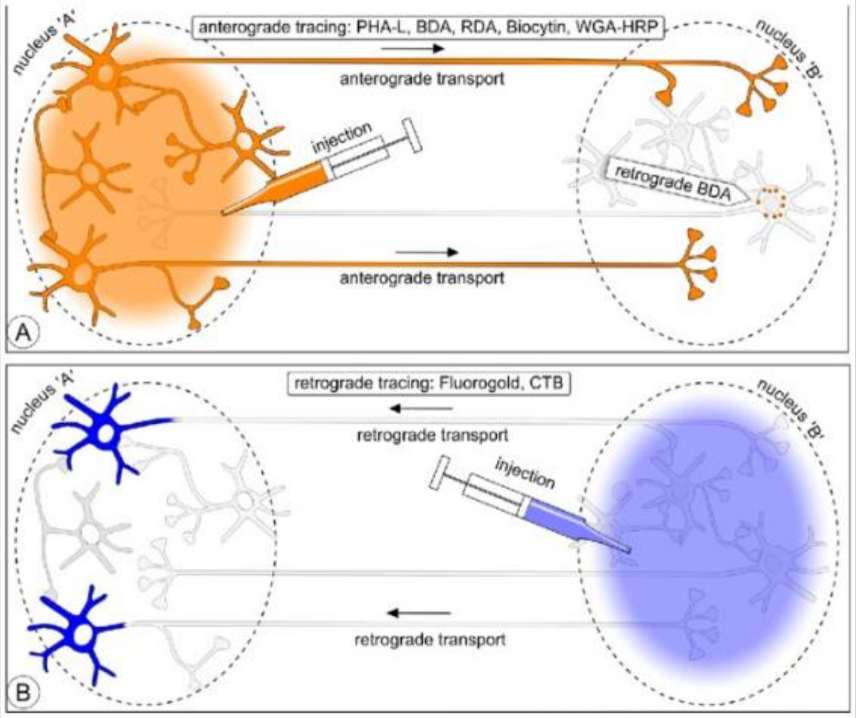
Balakrishnan, Guha, et al. "Voxelmorph: a learning framework for deformable medical image registration." *IEEE transactions on medical imaging* 38.8 (2019): 1788-1800.

The Importance of Registration in Neuroscience



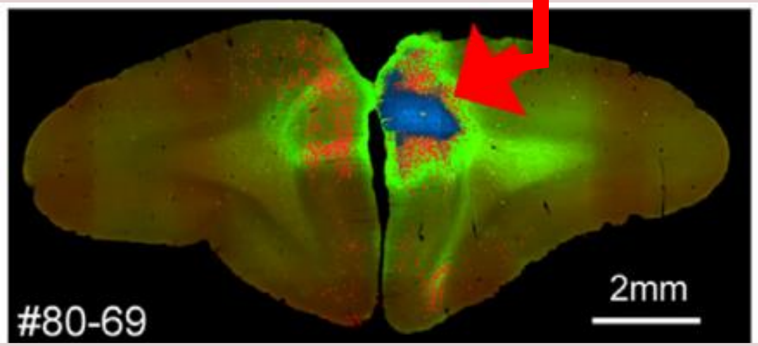


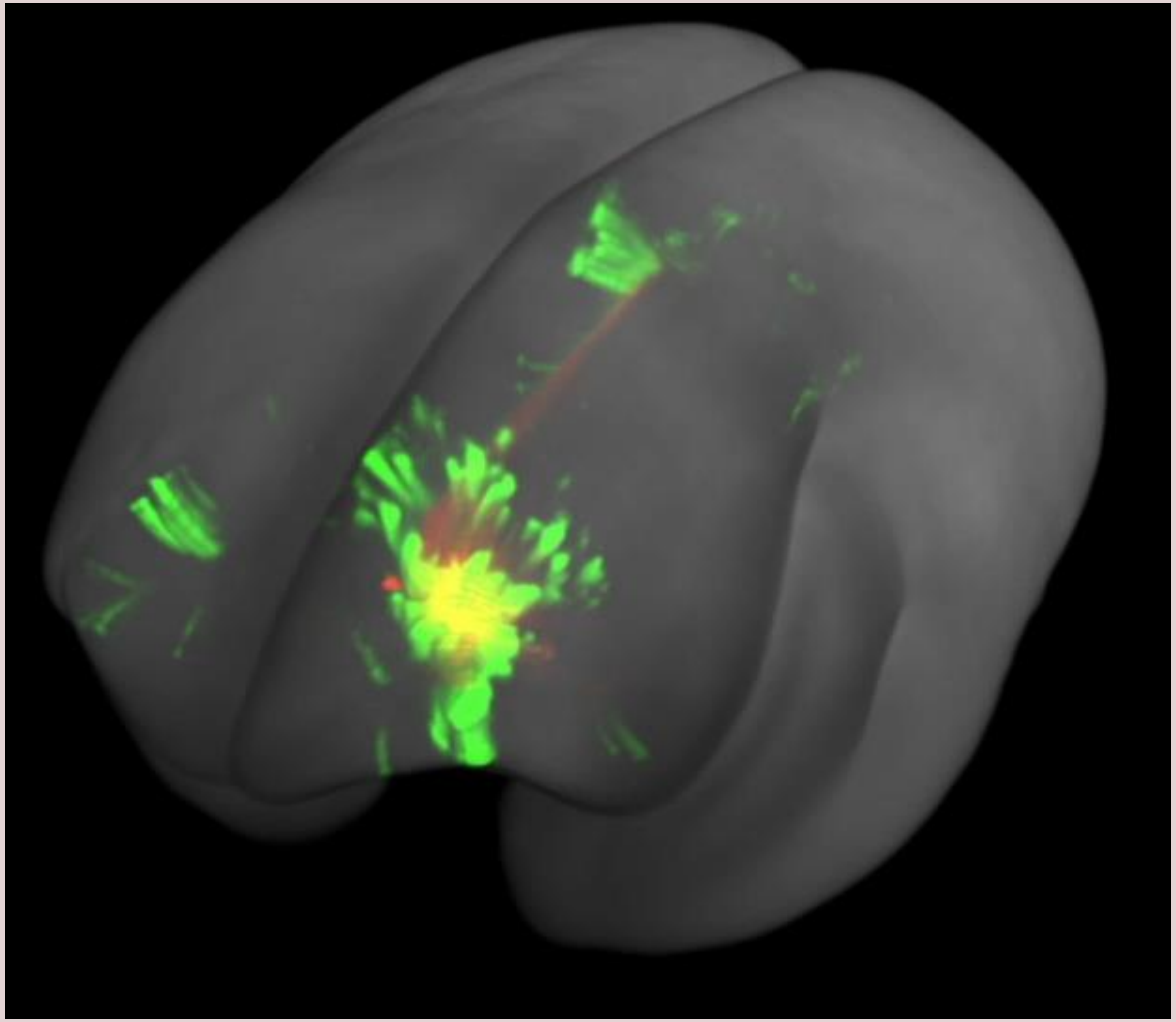
Neurons' Connectome Analysis of Marmoset Brain



Lanciego, J. L., & Wouterlood, F. G. (2020). Neuroanatomical tract-tracing techniques that did go viral. *Brain Structure and Function*, *225*(4), 1193-1224.

INJECTION SITE





Watakabe, Akiya, Henrik Skibbe, Ken Nakae, Hiroshi Abe, Noritaka Ichinohe, Muhammad Febrian Rachmadi, Jian Wang et al. "Local and long-distance organization of prefrontal cortex circuits in the marmoset brain." Neuron 111, no. 14 (2023): 2258-2273.

Explainable Artificial Intelligence (XAI) in Biomedical Image Analysis

Section 4

Case studies:

1. Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches.
2. Is grad-CAM explainable in medical images?
3. Grad-CAM: visual explanations from deep networks via gradient-based localization.

The Importance of XAI

- The **eXplainable AI (XAI)** is a new emerging research field because complex analysis tools, such as deep learning, are not easy for humans to understand (i.e., they are usually called black box methods).
- Generally, XAI refers to techniques that enable users/stakeholders to better understand an AI algorithm or system and its decisions.
- Explainability of a method is very important in the healthcare field, because of several reasons.
  1. To control and supervise AI systems by human experts where the consequences of such decisions can mean life and death.
  2. Interpretability and explainability enable the linking of clinicians' and doctors' domain expertise.
  3. Explainability not only aims at functional benefits but also adds to clinical confidence and consistent compliance with legal and ethical requirements (i.e., the final call must be made by a human expert).
  4. AI systems' interpretability and explainability enable human experts to identify errors, limitations, and potential biases in clinical settings.

Borys, Katarzyna, et al. "Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches." *European journal of radiology* 162 (2023): 110786.

Abstracted depiction of the interaction between AI models and end-users without XAI

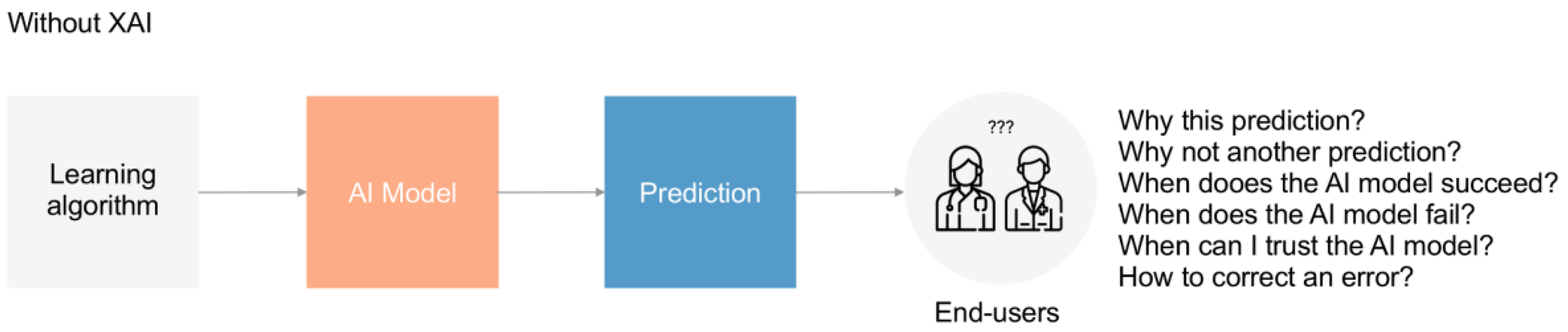


Fig. 1. Abstracted depiction of the interaction between AI models and end-users, which is not extended by the use of XAI.

Borys, Katarzyna, et al. "Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches." *European journal of radiology* 162 (2023): 110786.

Abstracted depiction of the interaction between AI models and end-users with XAI

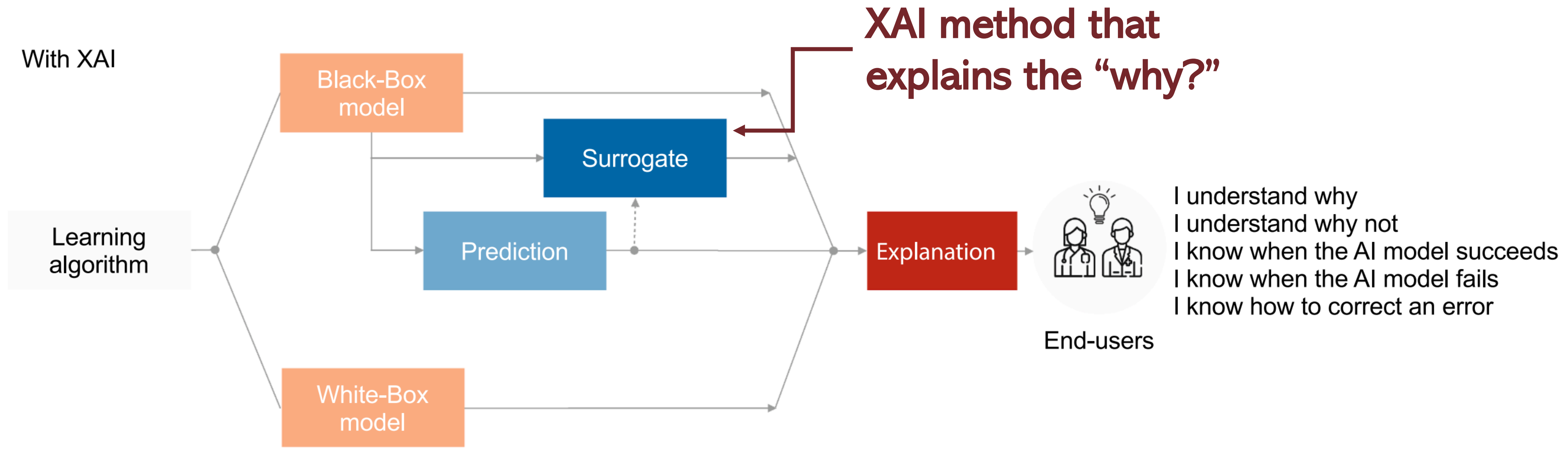
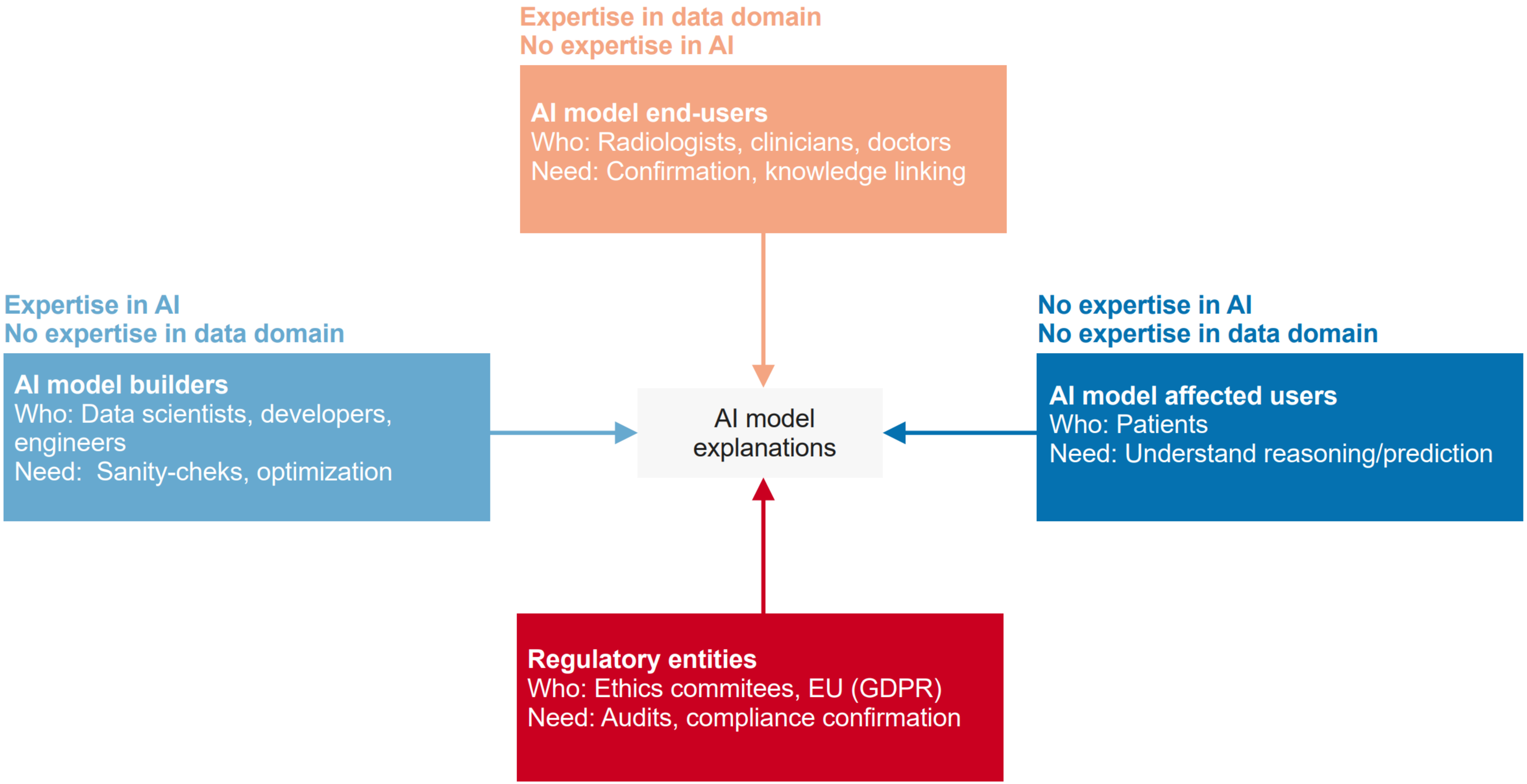


Fig. 2. Distinction of a deployment setup concerning an AI system with XAI. White-Box models are explainable by nature and can directly yield explanations. In contrast, Black-Box models must be explained globally (the model as a whole) or locally (a model's single prediction). Both views can optionally include the training of a subsequent surrogate model. Surrogate models are explainable models trained to approximate the predictions of an opaque Black-Box model.

Borys, Katarzyna, et al. "Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches." *European journal of radiology* 162 (2023): 110786.

The Importance of XAI in Healthcare



Borys, Katarzyna, et al. "Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches." *European journal of radiology* 162 (2023): 110786.

XAI Method Categorization

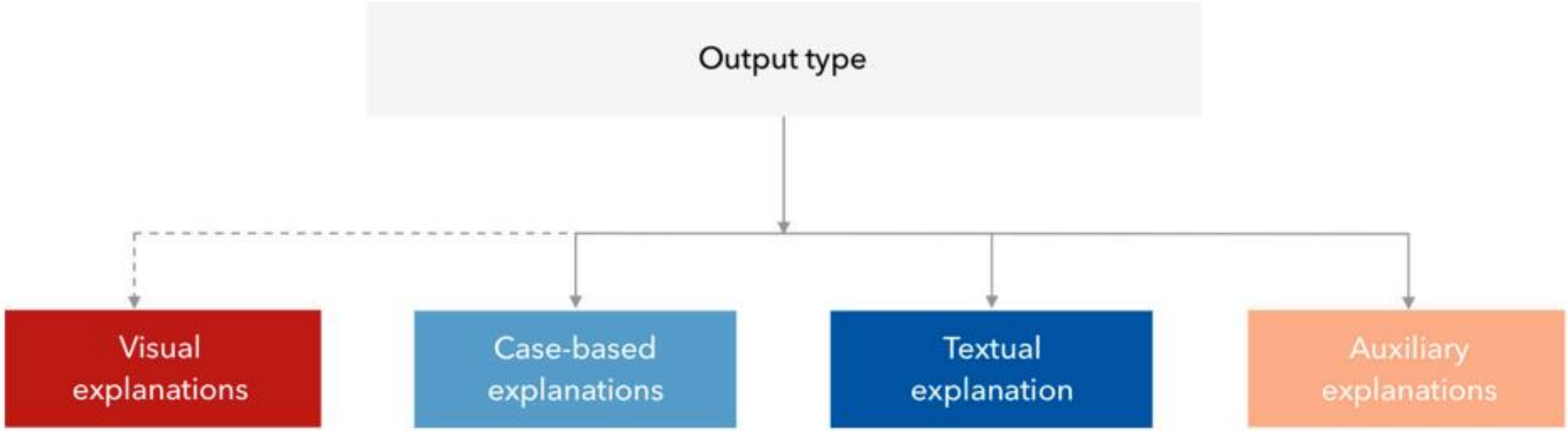


Fig. 1. Distinction of explanatory approaches by resulting output presentation form. Non-visual XAI approaches encompass auxiliary, case-based, and textual explanations. Visual explanations are not considered in this work but have been listed for completeness.

Borys, Katarzyna, et al. "Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches." *European journal of radiology* 162 (2023): 110786.

Visual XAI Methods in Biomedical Imaging

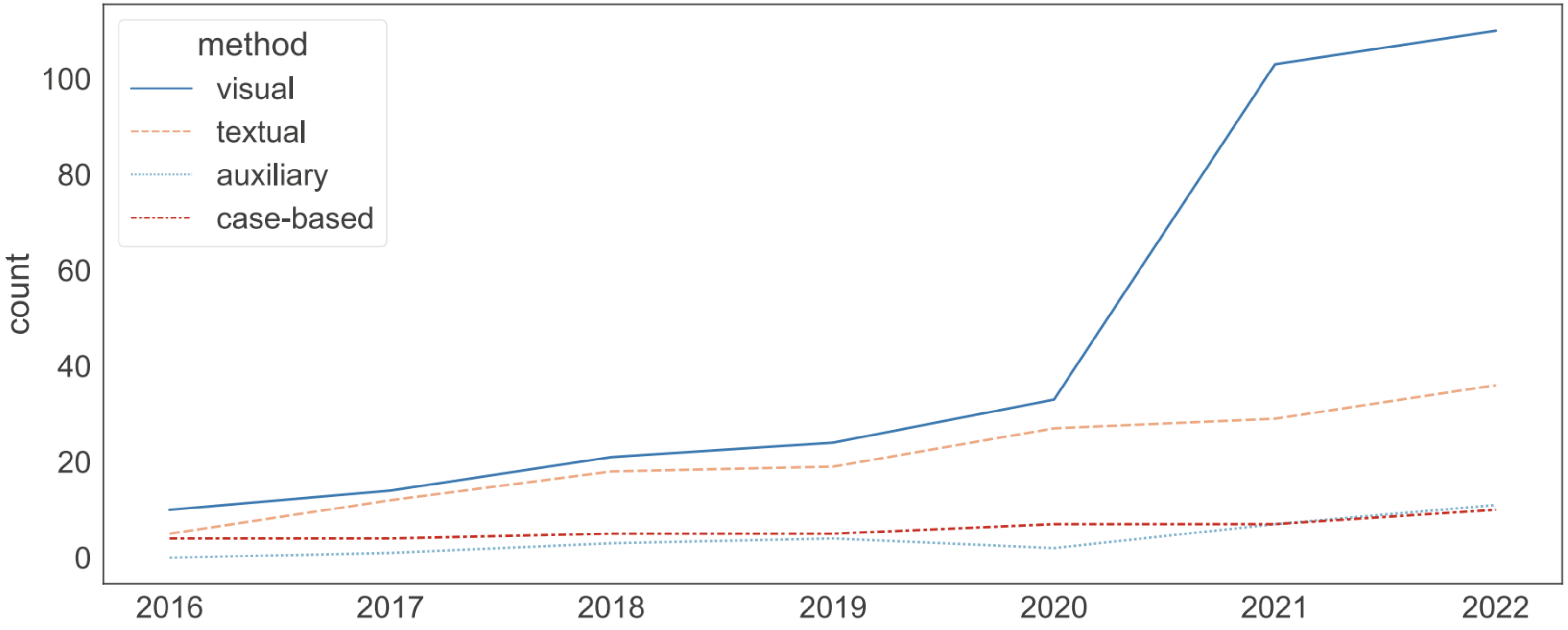


Fig. 5. Annual development of saliency-based (visual) and other non-visual (frequently divided into textual, example-based- case-based [21]) XAI methods applied in medical images, based on the cumulative number of citations (see Supplementary File 1).

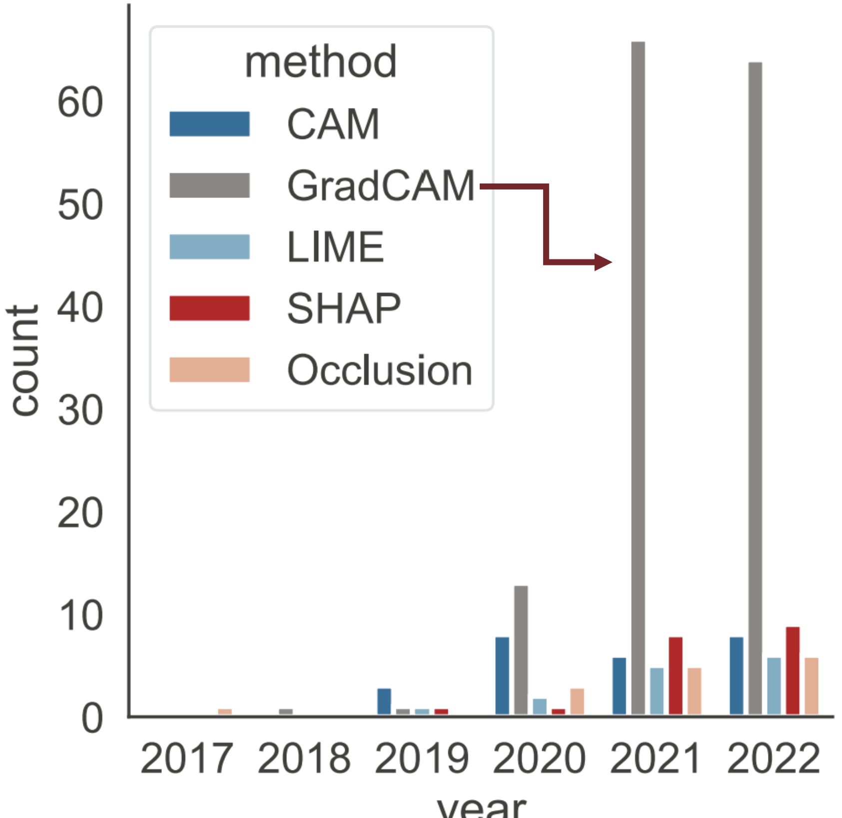


Fig. 6. Annual development of Top 5 saliency-based XAI methods applied in medical image analysis based on the total number of citations.

Borys, Katarzyna, et al. "Explainable AI in medical imaging: An overview for clinical practitioners–Beyond saliency-based XAI approaches." *European journal of radiology* 162 (2023): 110786.

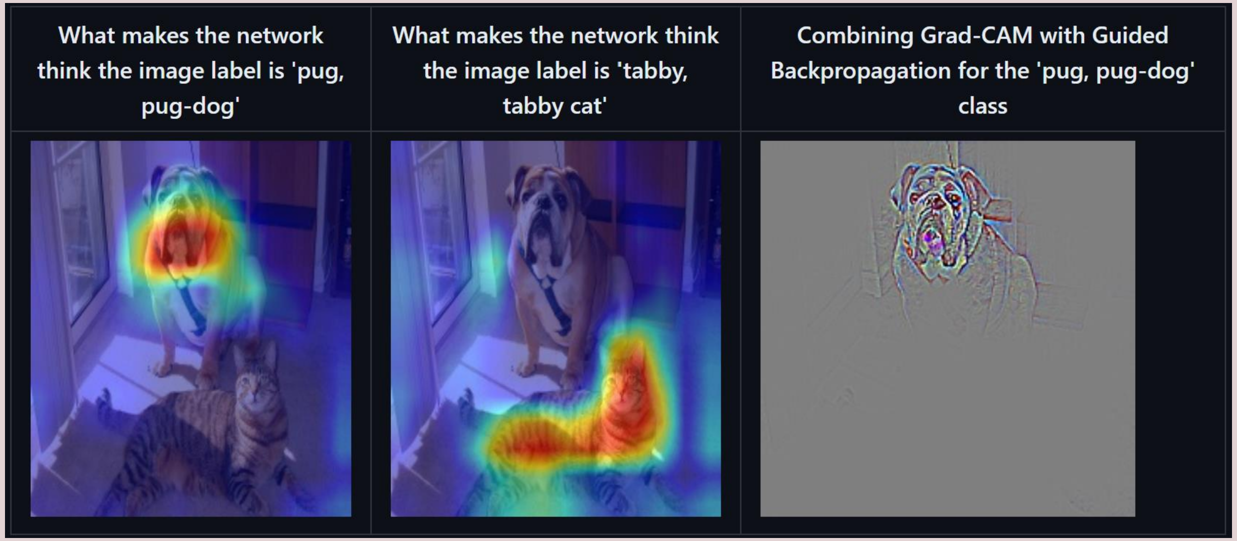
CAM-based Methods in Action

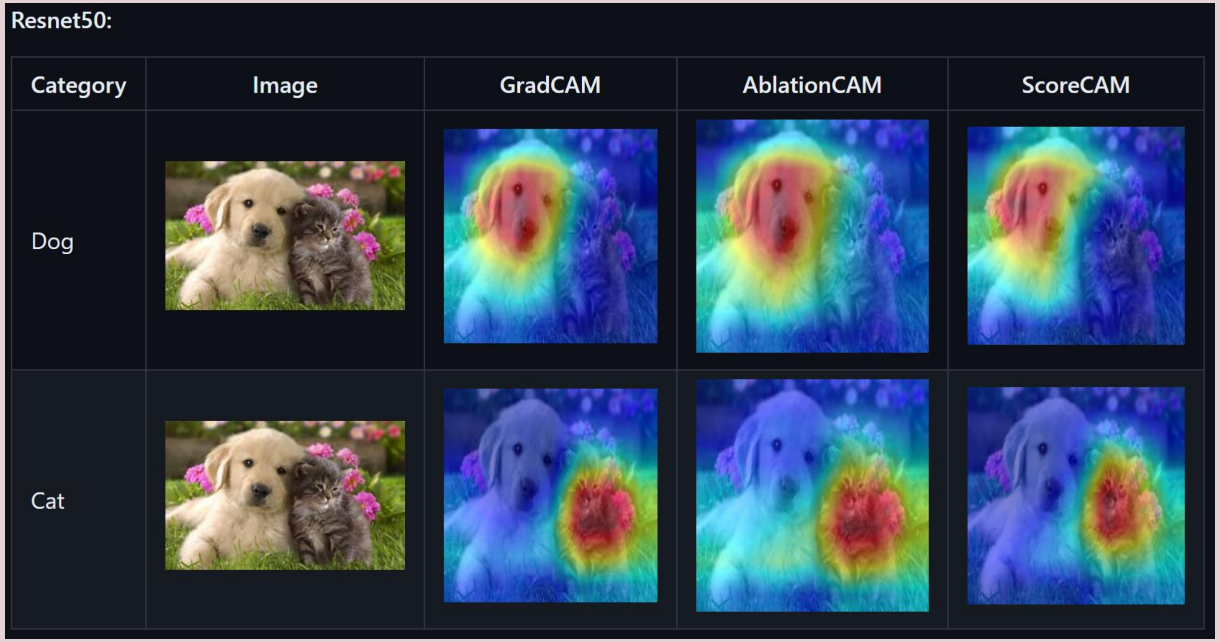
CAM = activation function

akan membuat mapping ke kelas-kelas tertentu

- Visual examples:

fokus ke gambar yang ingin kita cari, yang mau dicari bakal di highlight

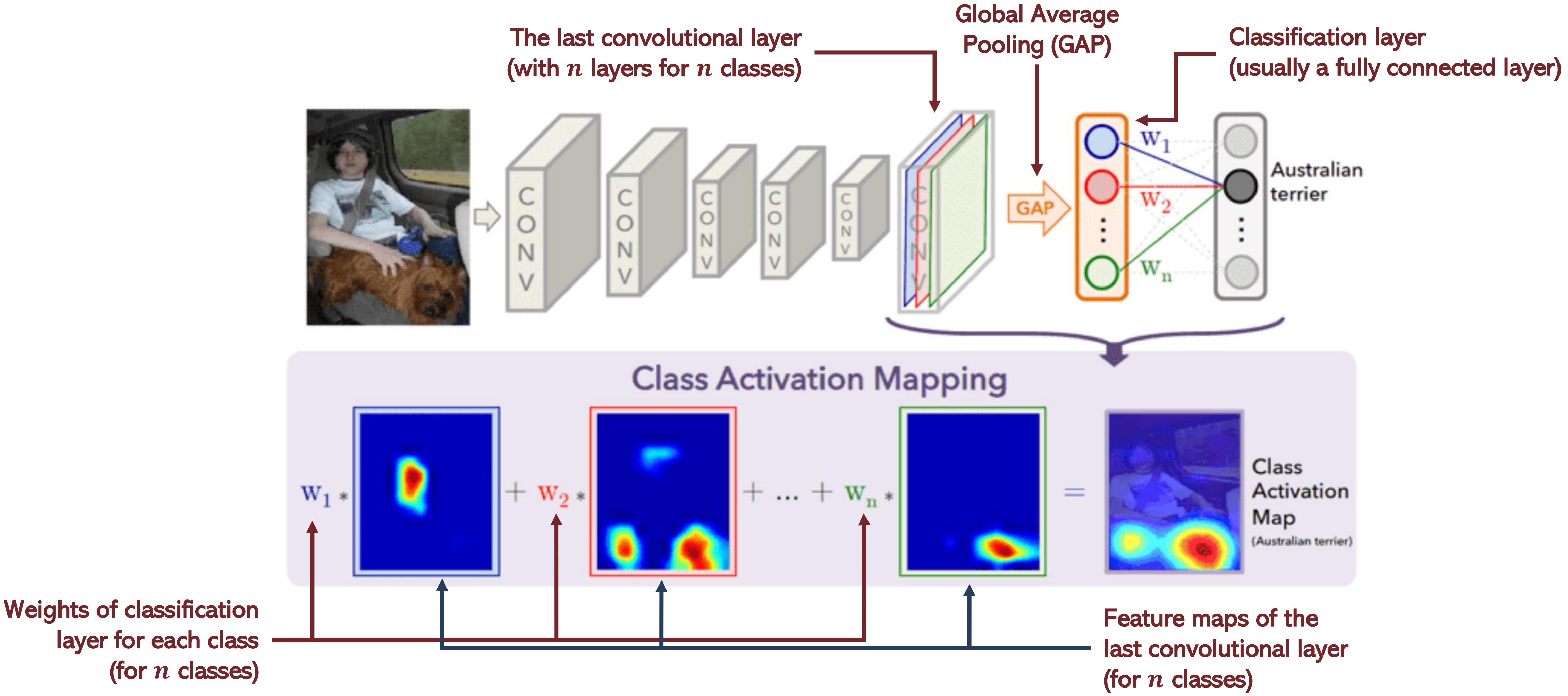




- Even works for videos!

https://github.com/jacobgil/pytorch-grad-cam

Grad-CAM
(Gradient-weighted Class Activation Mapping)



Suara, Subhashis, et al. "Is grad-CAM explainable in medical images?." *International Conference on Computer Vision and Image Processing*. Cham: Springer Nature Switzerland, 2023.

Grad-CAM’s Robustness

- Grad-CAM is a model- and task-agnostic XAI method.

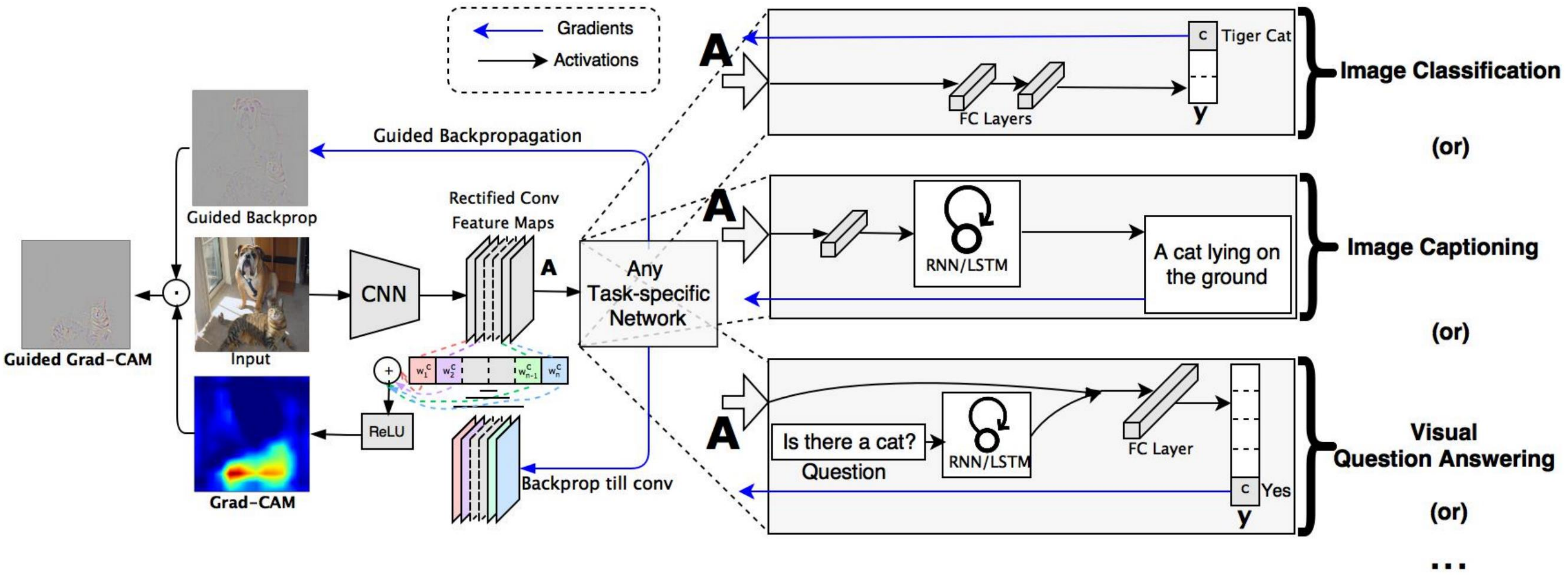
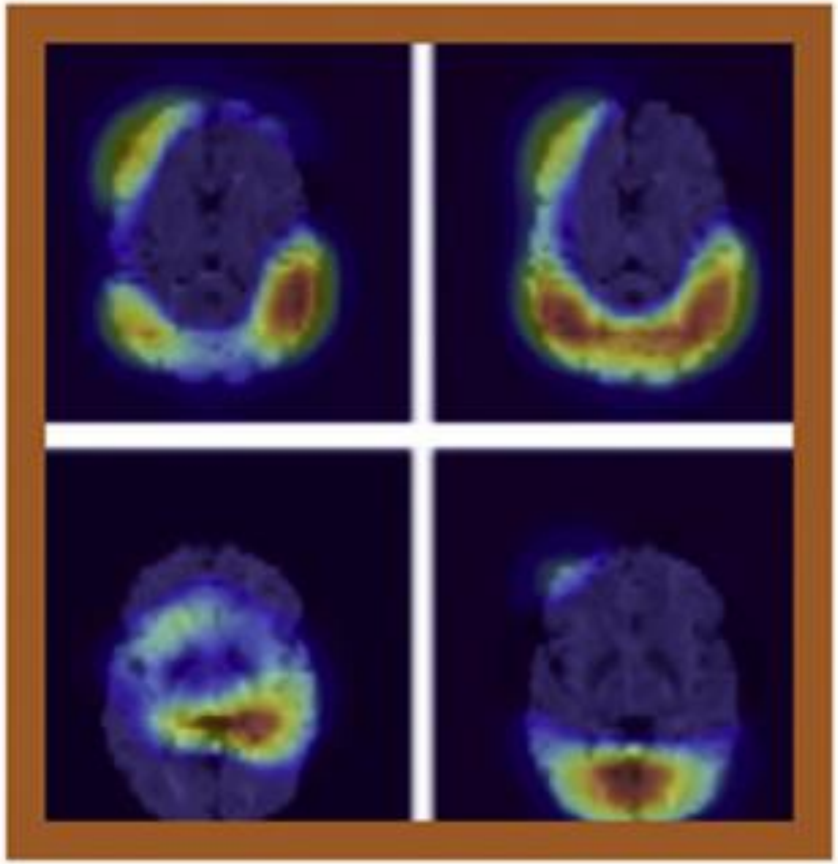


Fig. 2: Grad-CAM overview: Given an image and a class of interest (*e.g.*, ‘tiger cat’ or any other type of differentiable output) as input, we forward propagate the image through the CNN part of the model and then through task-specific computations to obtain a raw score for the category. The gradients are set to zero for all classes except the desired class (tiger cat), which is set to 1. This signal is then backpropagated to the rectified convolutional feature maps of interest, which we combine to compute the coarse Grad-CAM localization (blue heatmap) which represents where the model has to look to make the particular decision. Finally, we pointwise multiply the heatmap with guided backpropagation to get Guided Grad-CAM visualizations which are both high-resolution and concept-specific.

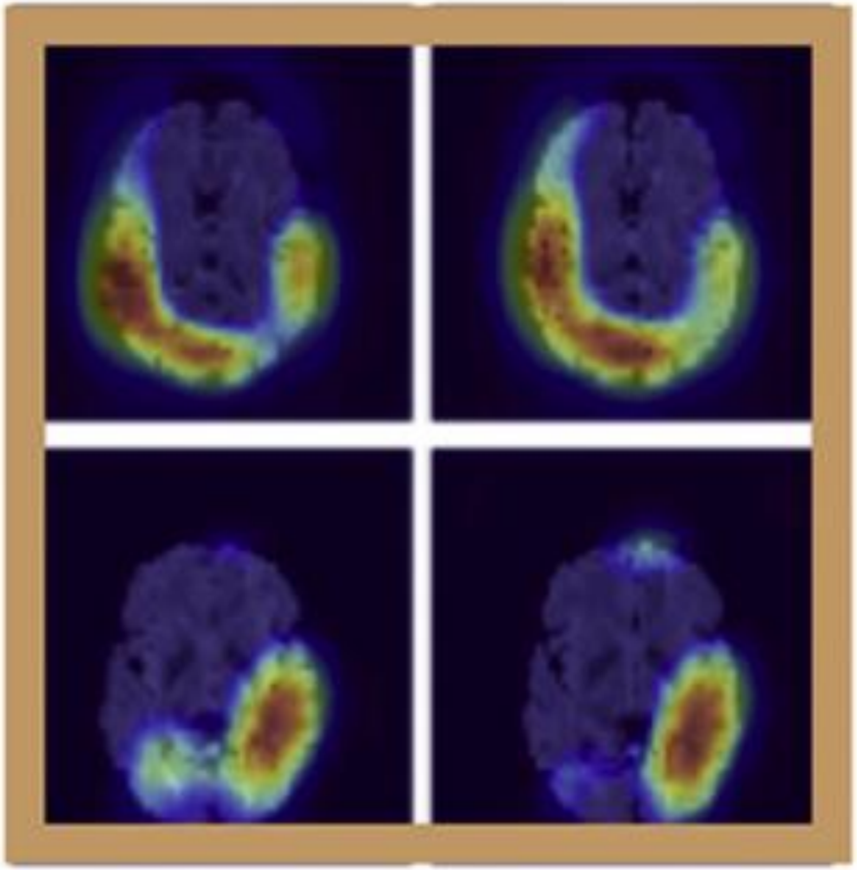
Selvaraju, Ramprasaath R., et al. "Grad-CAM: visual explanations from deep networks via gradient-based localization." *International journal of computer vision* 128 (2020): 336-359.

Grad-CAM to Classify Multiple Sclerosis Types

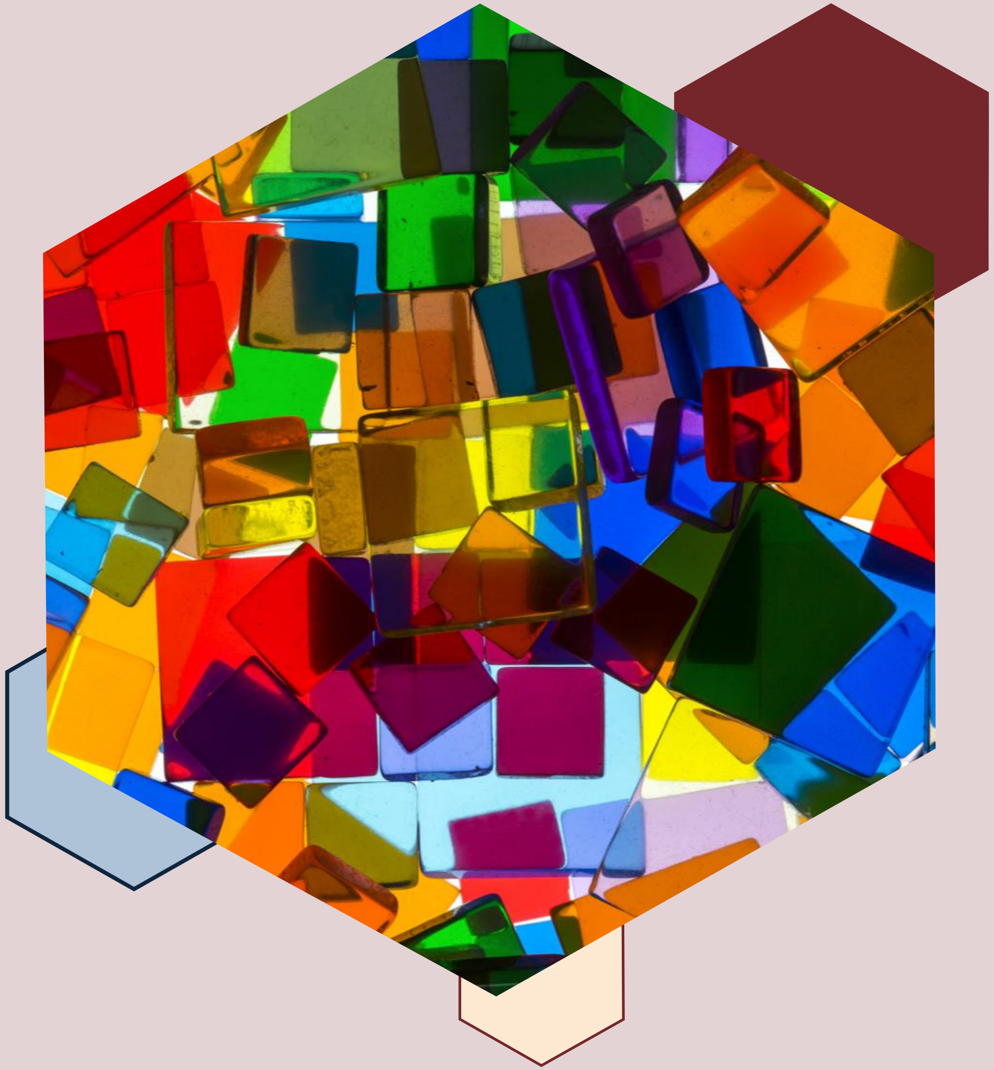
VGG16



VGG19



Zhang, Yunyan, et al. "Grad-CAM helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging." Journal of Neuroscience Methods 353 (2021): 109098.



Uncertainty in Biomedical Image Analysis

Section 5

Case study: A Probabilistic U-Net for Segmentation of Ambiguous Images

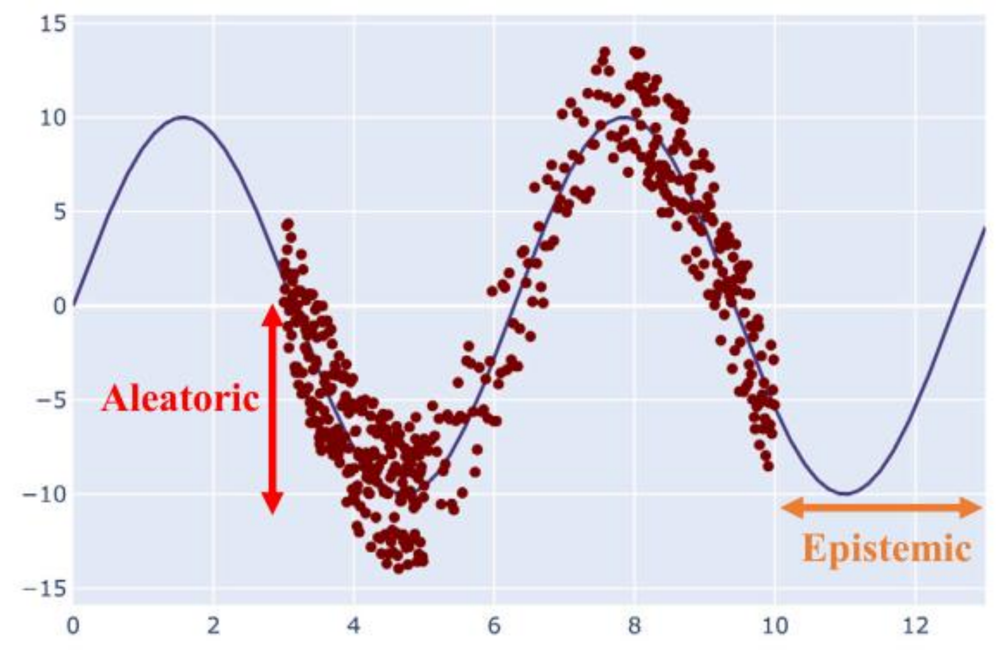
Uncertainty in Machine Learning

- There are two different types of uncertainty in machine learning.
  - **Aleatoric (statistical) uncertainty** refers to the notion of randomness that comes from the data itself (i.e., the variability in the outcome of an experiment, which is due to inherent noises of the data).
    - This uncertainty **cannot** be solved by adding more data.
  - **Epistemic (systematic) uncertainty** refers to uncertainty caused by a lack of knowledge (data) or ignorance about the best model (underfitting/overfitting).
    - This uncertainty (usually) **can** be solved by adding more data.

Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine Learning*, *110*(3), 457-506.

Aleatoric vs. Epistemic

- **Aleatoric uncertainty** is usually high in overlapping regions (near decision boundaries).
- **Epistemic uncertainty** is usually high in regions where there is a lack of training examples.



Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., ... & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion*, *76*, 243-297.

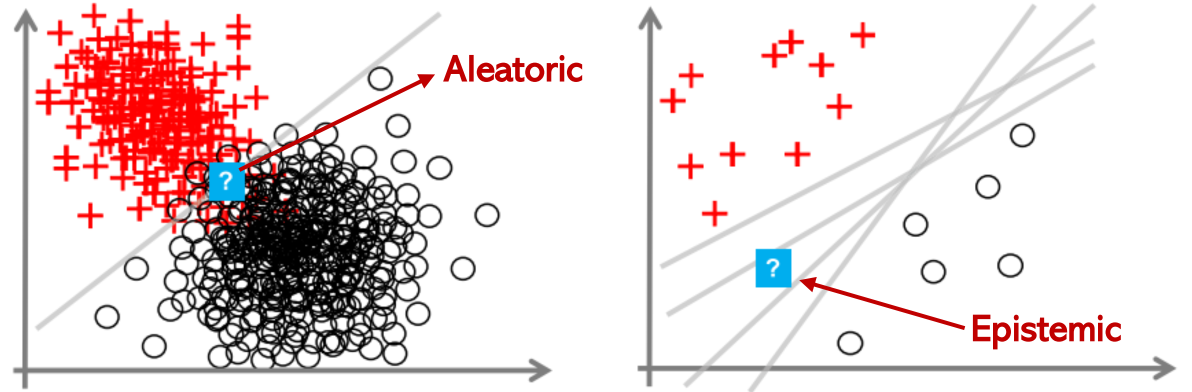


Fig. 6 Left: Even with precise knowledge about the optimal hypothesis, the prediction at the query point (indicated by a question mark) is aleatorically uncertain, because the two classes are overlapping in that region. Right: A case of epistemic uncertainty due to a lack of knowledge about the right hypothesis, which is in turn caused by a lack of data

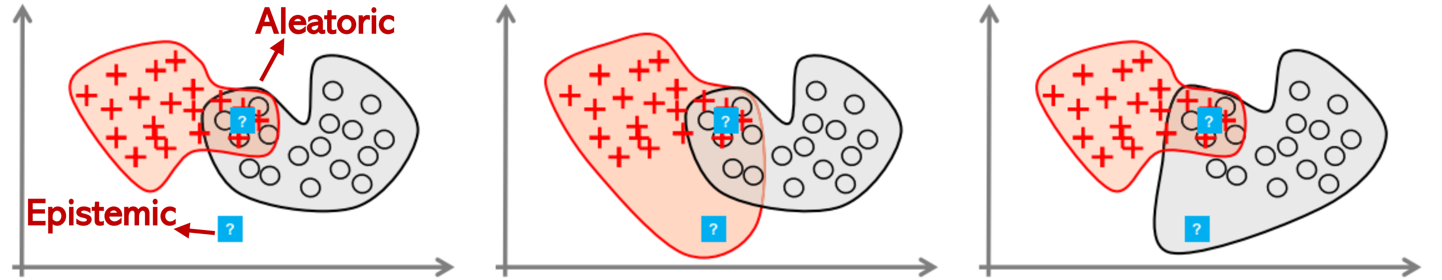
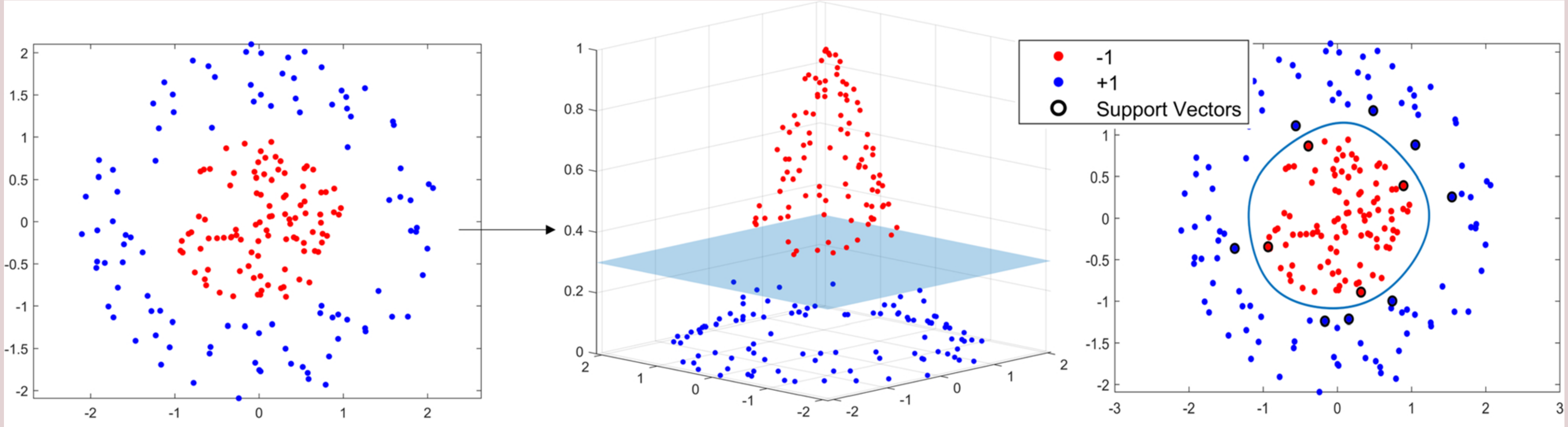


Fig. 12 If a hypothesis space is very rich, so that it can essentially produce every decision boundary, the epistemic uncertainty is high in sparsely populated regions without training examples. The lower query point in the left picture, for example, could easily be classified as red (middle picture) but also as black (right picture). In contrast to this, the second query is in a region where the two class distributions are overlapping, and where the aleatoric uncertainty is high

Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine Learning*, *110*(3), 457-506.

What to do when facing aleatoric uncertainty?

- Identify whether you are facing overfitting or underfitting.
  - If overfitting, try to simplify the problem. Remember cascade classification?
  - If underfitting, you might have to use a more complex method.



(a) 2D non-linear training data

(b) 3D mapped data points using the Gaussian kernel and separating hyperplane

(c) non-linear SVM with the Gaussian kernel training result

- Try to understand your data better. Try to visualize it, if possible.
- Try to change your assumption. Machine learning is always about assumption.

Wang, Yanran, et al. "Support vector machine guided reproducing kernel particle method for image-based modeling of microstructures." *Computational Mechanics* 73.4 (2024): 907-942.

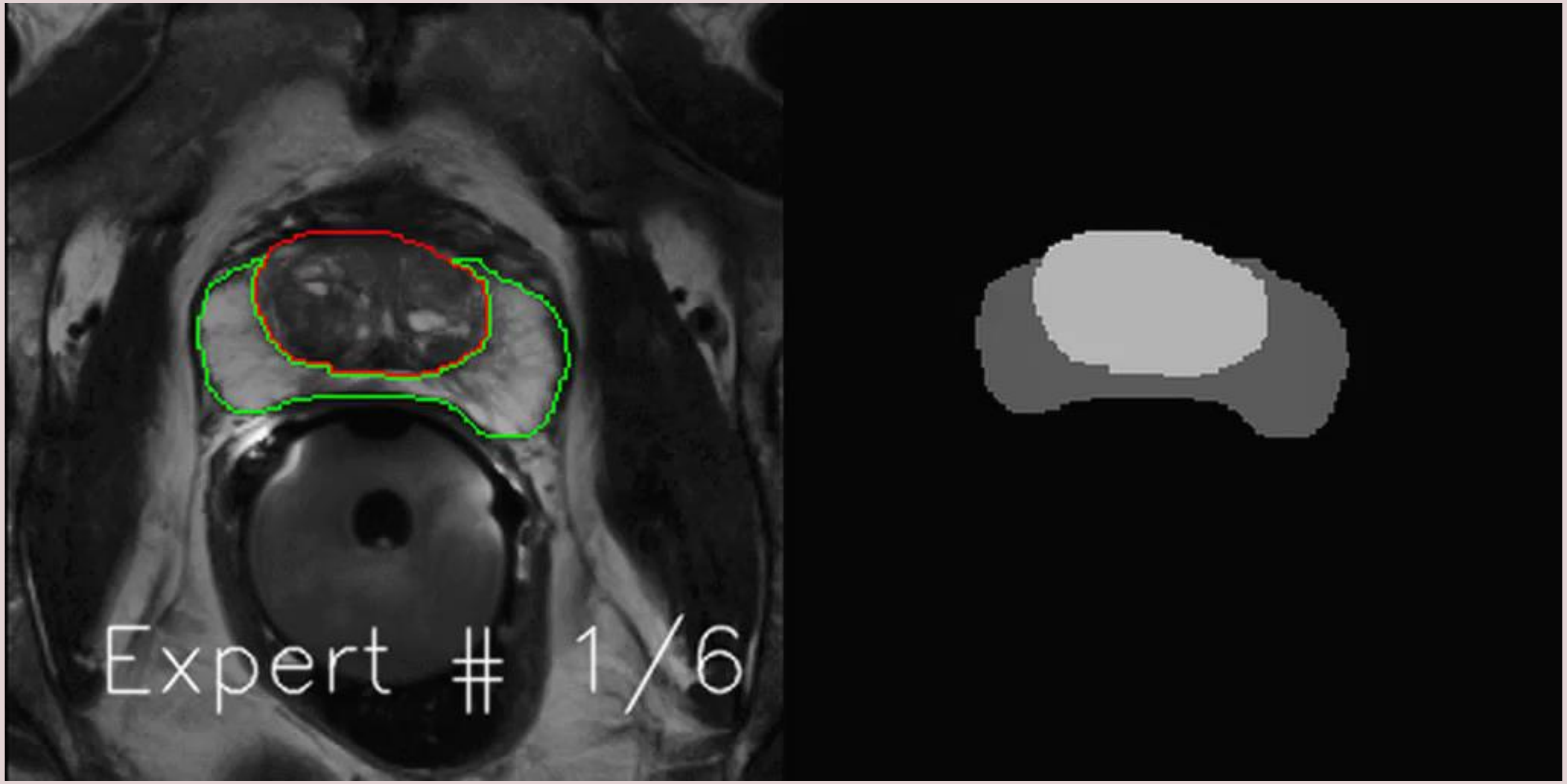
Uncertainty Quantification (UQ)

- The two most popular uncertainty quantification (UQ) methods in ML:
  - Bayesian techniques
  - Ensemble techniques
- **Bayesian techniques**: Bayesian methods are robust to overfitting problems and can be trained on both small and large (big) datasets. However, many Bayesian methods are not computationally effective (expensive).
  - Examples: Monte Carlo (MC) dropout, Markov chain Monte Carlo (MCMC), variational inference (VI), variational autoencoders, etc.
- **Ensemble techniques**: These involve using multiple machine learning models trained in different datasets for better generalization. They are expensive, but multiple studies have shown that this approach is effective in unseen test data.
- Please take the machine learning class if you are interested in these concepts!

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., ... & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion*, *76*, 243-297.

Aleatoric Uncertainty in Biomedical Image Analysis

1. **Inter-rater Reliability:** Tingkat kesepakatan (agreement) dari berbagai pakar penilai (expert raters) terhadap sebuah (kumpulan) data.
2. **Intra-rater Reliability:** Tingkat kesepakatan (agreement) dari beberapa penilaian yang dilakukan oleh seorang pakar penilai (expert rater) terhadap sebuah (kumpulan) data.



Christian F. Baumgartner (@c_f_baumgartner). Link: https://twitter.com/c_f_baumgartner/status/1138457478763622402

Probabilistic U-Net

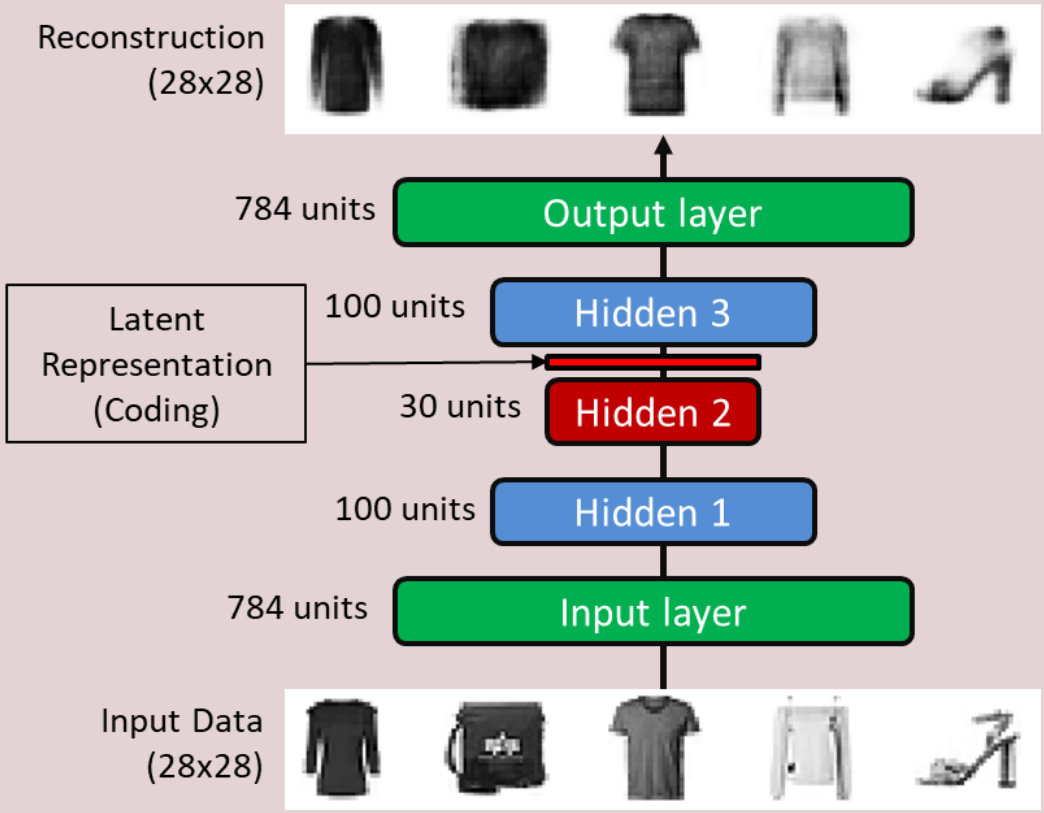
- Combination of **variational autoencoder** (VAE) and U-Net.
- **Variational autoencoder (VAE)** is a little bit different from the other autoencoders:
  1. VAE is a probabilistic autoencoder where the output of the model is determined randomly using the sampling method from a latent space that is noised at the time of prediction (after training).
  2. VAE is a generative autoencoder where VAE generates a new instance that looks like it was taken from the training set.

Kohl, Simon, et al. "A probabilistic u-net for segmentation of ambiguous images." Advances in Neural Information Processing Systems. 2018.

U-Net vs. VAE

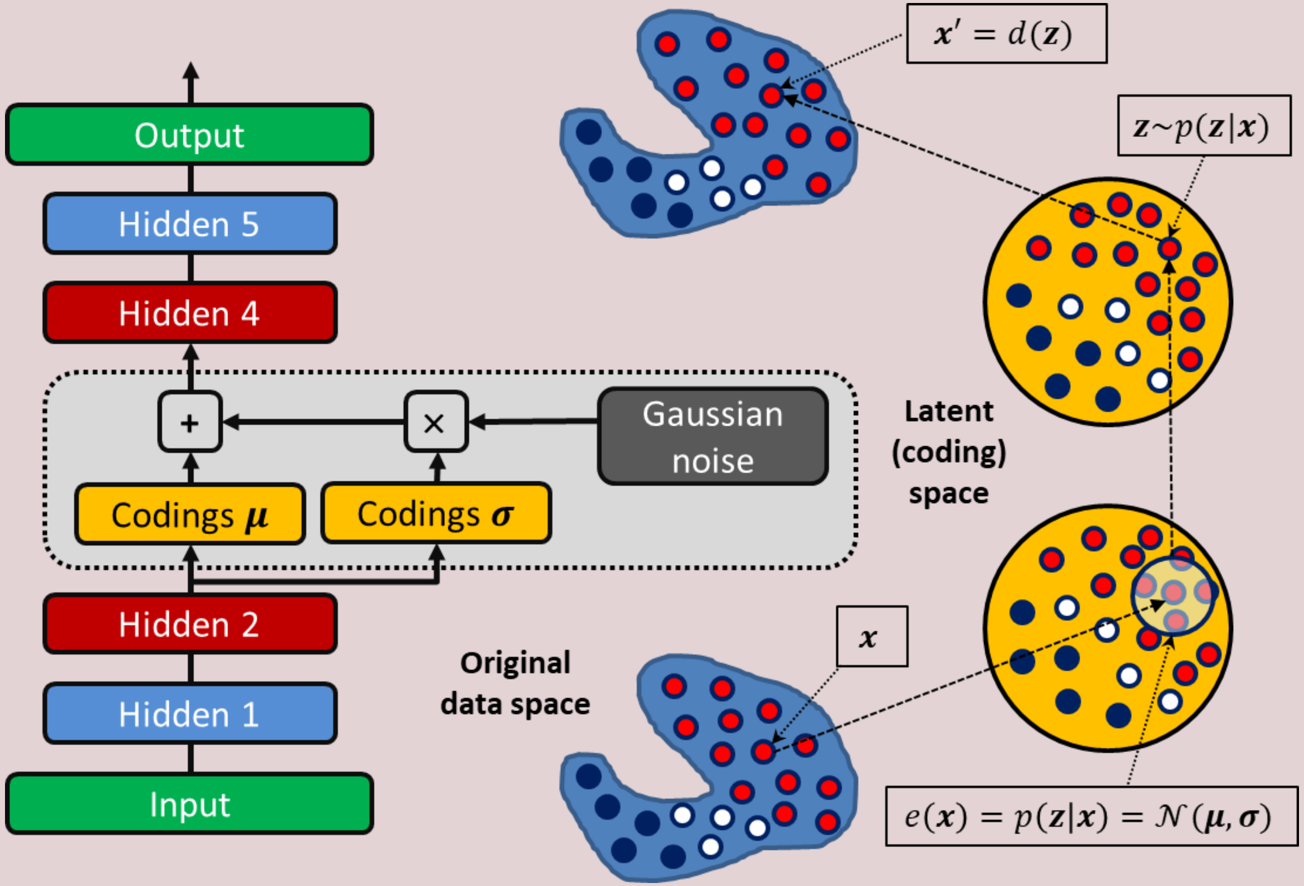
U-Net

- The latent representation (coding) is a set of feature maps.



VAE

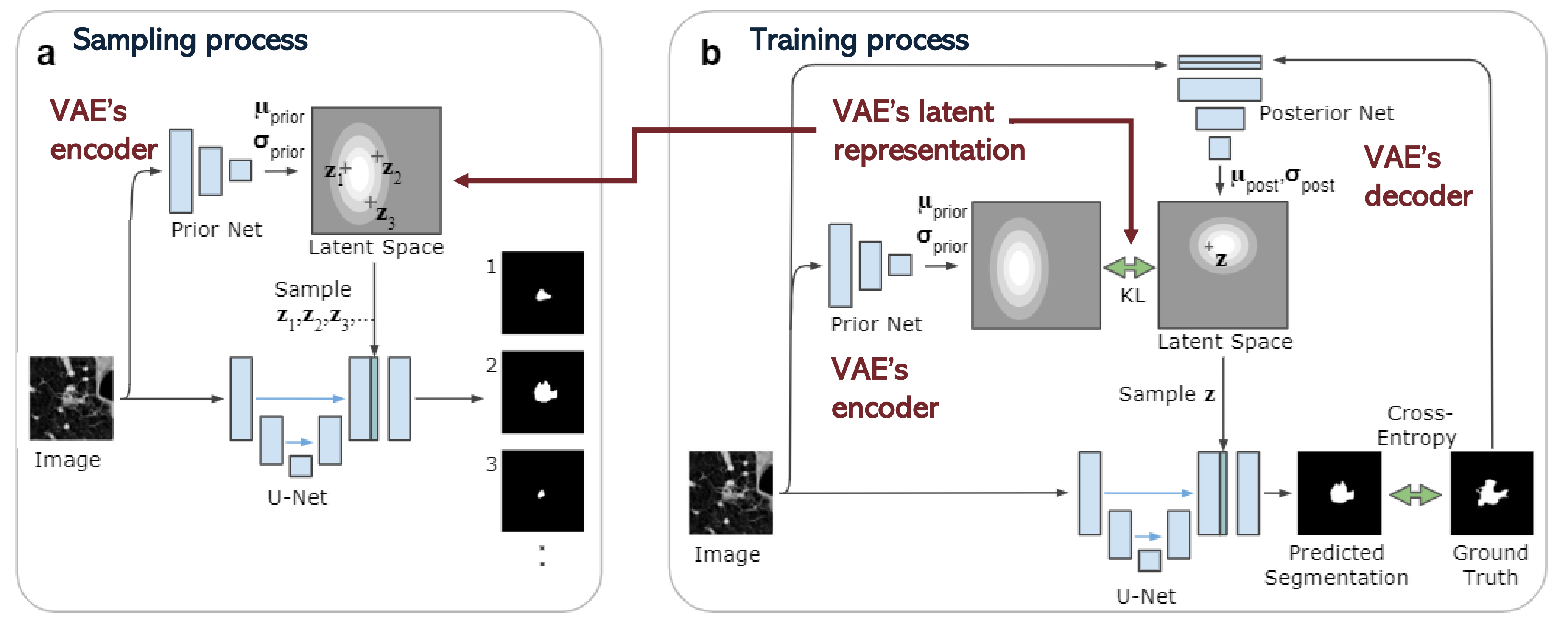
- The latent representation (coding) is a probabilistic distribution (e.g., Gaussian).



Architecture of the Probabilistic U-Net

a) **Sampling process** (<u>pada saat prediksi</u>). The heatmap represents the probability distribution in the low-dimensional latent space RN (e.g., N = 6 in our experiments). For each network execution, one sample z → RN is drawn to predict one segmentation mask.

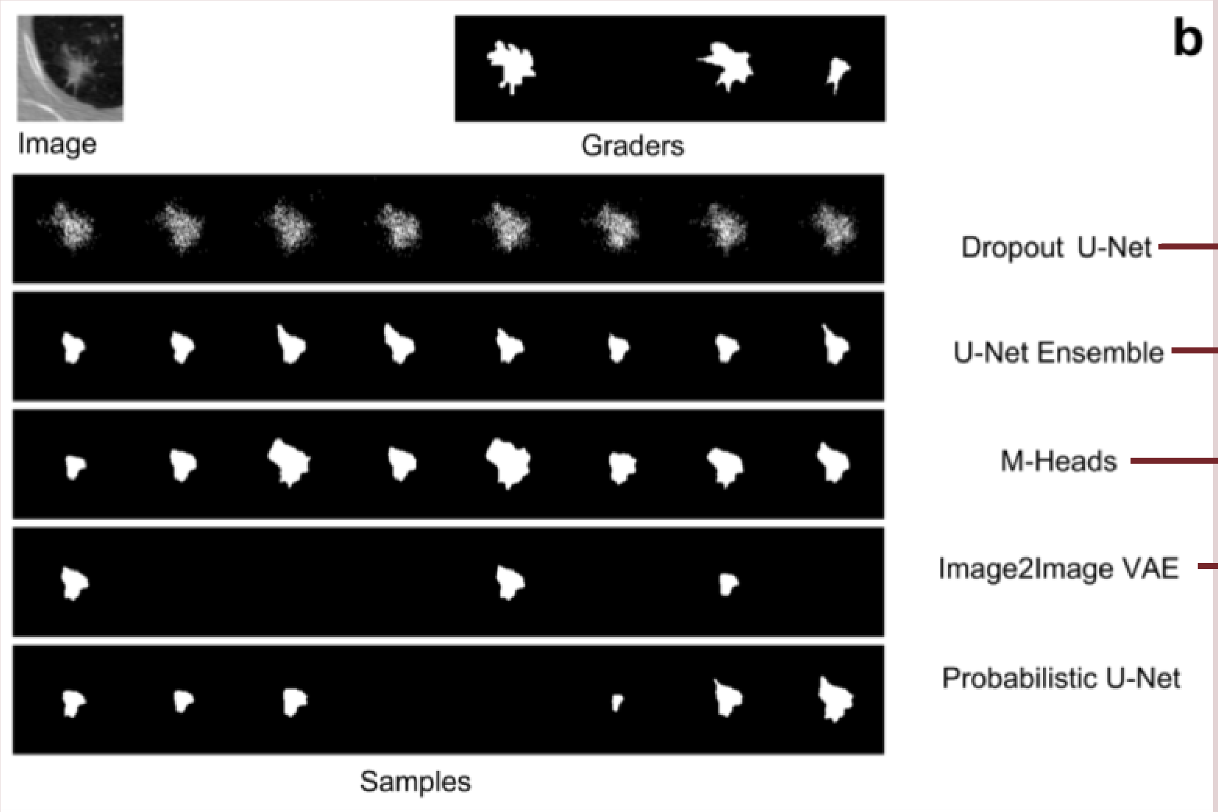
b) **Training process** illustrated for one training example. *Green arrows*: loss functions.



Kohl, Simon, et al. "A probabilistic u-net for segmentation of ambiguous images." Advances in Neural Information Processing Systems. 2018.

Results From the Probabilistic U-Net

- Lung CT scan from the LIDC test set. Ground truth: 4 graders.



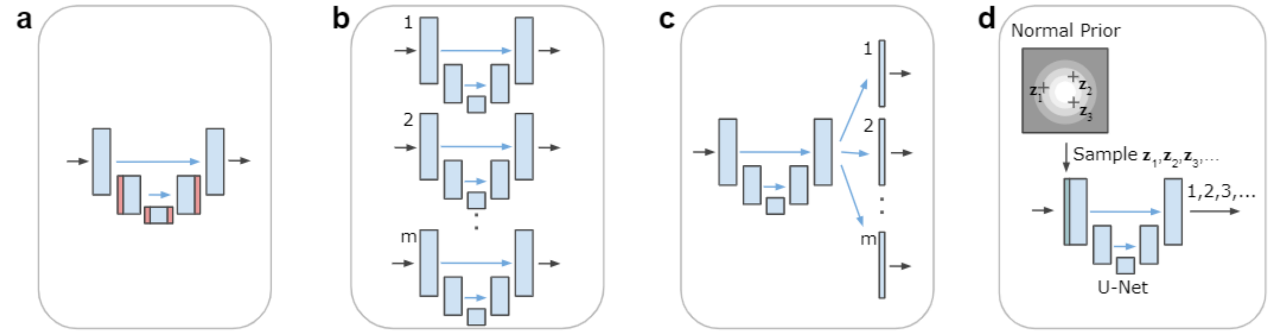


Figure 2: Baseline architectures. Arrows: flow of operations; blue blocks: feature maps; red blocks: feature maps with dropout; green block broadcasted latents. Note that the number of feature map blocks shown is reduced for clarity of presentation. (**a**) Dropout U-Net. (**b**) U-Net Ensemble. (**c**) M-Heads. (**d**) Image2Image VAE.

Kohl, Simon, et al. "A probabilistic u-net for segmentation of ambiguous images." Advances in Neural Information Processing Systems. 2018.

Thank you!

CSCE604133 Computer Vision
Faculty of Computer Science
Universitas Indonesia

Dr. Eng. Laksmita Rahadianti
Muhammad Febrian Rachmadi, Ph.D.
Dr. Dina Chahyati, Prof. Dr. Aniati M. Arymurthy