Assignment 3

Problem 1: Imaging principles of X-ray and CT

X-rays are a form of high-energy electromagnetic radiation that can pass through most objects including the body. In X-ray imaging, X-rays are directed towards the body and as they pass through different tissues, they are absorbed to varying degrees based on tissue density and atomic composition. A detector on the opposite side of the body captures the transmitted X-rays, converting them into an electrical signal that is processed to generate an image. [1]

Filtered back projection is a reconstruction technique used in computed tomography (CT) imaging. It mainly consists of two parts: filtration and projection. In the first process, the acquired projections are filtered to remove artefacts and enhance image quality. Then the filtered projections are then back-projected onto a two-dimensional image plane, and the accumulation of these back-projections creates the final reconstructed image, providing detailed information about the internal structures of the body in the imaged slice. The filtered back projection allows for visualisation and analysis of anatomical structures with high spatial resolution. [2]

X-ray imaging and CT are closely related techniques that utilise X-rays to visualise internal structures of the body. CT imaging combines multiple X-ray projections taken from different angles around the body. This is done by having the X-ray source and detector rotate around the person. Then, the data is processed through the filtered back projection mentioned above to reconstruct the collected projection data into meaningful cross-sectional images. [3]

Problem 2: Eye anatomical structures and eye diseases fundamentals

Humans perceive objects via eye structures. It begins when light enters the cornea which is the transparent outer covering of the eye and helps focus on incoming light. It then passes through the pupil in the centre of the iris, which controls the amount of light entering the eye. Behind the pupil is the lens which helps focus the light onto the retina. The retina, located at the back of the eye, contains photoreceptors cells which detect light and convert it into electrical signals. Once the photoreceptor cells capture light and convert it into electrical signals, these signals are transmitted to the optic nerve. The optic nerve carries the electrical signals from the retina to the brain where the signals are processed and interpreted for humans to perceive objects. [4]

One ophthalmology disease is the Diabetic retinopathy, which occurs due to prolonged high blood sugar levels that damage the blood vessels supplying the retina. The disruption of normal blood flow and the leakage of fluid and blood interfere with the function of the retinal photoreceptor cells, causing visual abnormalities. [5]

One imaging technique used to detect ophthalmology diseases is through the Fundus camera. A Fundus camera captures images of the fundus, which is the interior surface of the eye including the retina, optic disc, and blood vessels. The images captured by the Fundus camera can help detect the diseases as the images can reveal the progression of certain eye conditions and diseases. Fundus photography is commonly used to detect and monitor diseases such as diabetic retinopathy, macular degeneration, glaucoma, and retinal vascular disorders. [6] [7]

Another imaging technique used to detect ophthalmology diseases is through optical coherence tomography. Optical coherence tomography generates cross-sectional images of the eye by using low-coherence light to create high-resolution, real-time images of the retinal layers. It works on the principle of interferometry, measuring the echo time delay and intensity of light reflected from different tissue layers. This allows for the visualisation of the retinal thickness, macular features, and the integrity of various structures, which is useful for detecting diseases. Optical coherence tomography can be used in detecting conditions such as macular edema, macular holes, retinal detachment, and glaucoma. [6]

In relation to Diabetic retinopathy, optical coherence tomography can be used to detect Diabetic retinopathy. For example, the images taken from optical coherence tomography can be used to measure the swelling and thickness caused by Diabetic retinopathy in the retina. Another example is that optical coherence tomography can be used to detect the fluid accumulation caused by the leakage of fluid from damaged blood vessels due to diabetic retinopathy. [8]

Problem 3: Unsupervised Learning

Supervised learning involves training a model on a labelled dataset, where each data instance is associated with a corresponding target or output label. The model learns from the provided labelled examples and generalises to make predictions on new unlabeled instances. [10]

In contrast, unsupervised learning deals with unlabeled data, where the input instances do not have corresponding target labels. Whereas supervised learning is more focused on finding the relationship between input and output data, unsupervised learning is more focused on discovering patterns and relationships within the unlabelled data. In summary, supervised learning focuses on prediction accuracy, while unsupervised learning emphasises data exploration and understanding. [9]

Autoencoder is a neural network that uses the encoder-decoder architecture. The encoder compresses the input data into a lower-dimensional representation known as a latent space. The decoder then reconstructs the original input from this latent space. The autoencoder tries to minimise the loss of information (difference between the input and the reconstructed output) measured by a loss function. [10]

Variational Autoencoder (VAE) is similar to an autoencoder and also uses the encoder-decoder architecture. However, instead of directly encoding the input into a deterministic latent code, VAE models the latent space as a probability distribution. The encoder generates the parameters (mean and variance) of this distribution, and the decoder samples from it to reconstruct the input. [10]

Given a set of fundus photos without labels, a representation learning pipeline can be established using contrastive learning. In contrastive learning, the technique aims to learn useful representations by contrasting similar and dissimilar pairs of data samples. A contrastive loss function is defined to measure the similarity or dissimilarity between pairs of samples. These pairs are passed through a model which learns to map the input data into a lower-dimensional feature space. [11]

So to implement the pipeline above, first the fundus images are preprocessed to enhance the various features of the data, which can include resizing images and applying augmentation.

Then, an encoder architecture needs to be implemented. For fundus photos, using CNNs such as UNet would be more appropriate as the encoder for this task due to their ability to capture spatial patterns in images. We need to define a loss function for contrastive learning to differentiate between similar and dissimilar samples. One loss function that can be used is the Noise Contrastive Estimation (NCE), where the idea is to run logistic regression to tell apart the target data from noise. We then train the encoder by iteratively optimizing the loss function and updating the parameters of the encoder network, which can then be evaluated with validation and testing dataset. [12] [13]

Problem 4: The basic principles of Pathology.

One of the difficulties in applying AI models to Whole Slide Images (WSIs) is the large size of these images. WSIs are extremely large digital images that can range anywhere from gigabytes to terabytes in size [15]. This means that processing and analyzing WSIs require significant computational resources as the AI models need to handle massive amounts of data which can strain memory and processing capabilities.

To use WSIs as inputs for deep learning methods, most approaches do not use the whole image but rather extract and use only a small number of patches. The method of annotating WSIs can vary, with the two common methods being patch level annotation and slide level annotation. In patch level annotation, experts or annotators manually mark or identify specific regions of interest within each patch whereas in slide level annotation it involves annotating the entire WSI as a whole. Slide level annotations are more common due to practical limitations. [14]

Once the patches are extracted and annotated, preprocessing techniques such as colour normalisation can be done to reduce the number of patches that need to be analysed. During the training phase, the adjusted patches, along with their corresponding labels, are used to train the deep learning model. In the test stage, the adjusted patches are fed into the trained model to obtain predictions or feature representations, which can be further analysed for various tasks such as classification, segmentation, or object detection in WSIs. [15]

References

[1] “X-rays,” National Institute of Biomedical Imaging and Bioengineering, https://www.nibib.nih.gov/science-education/science-topics/x-rays (accessed Apr. 13, 2024).

[2] B. Nett, “Filtered BackProjection (FBP) Illustrated Guide for Radiologic Technologists,” How Radiology Works, https://howradiologyworks.com/filtered-backprojection-fbp-illustrated-guide-for-radiologic-technologists/ (accessed Apr. 13, 2024).

[3] L. M. Fayad, “CT scan versus MRI versus X-ray: What type of imaging do I need?,” Johns Hopkins Medicine, https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/ct-vs-mri-vs-xray (accessed Apr. 14, 2024).

[4] C. Stangor and J. Walinga, “Seeing,” Introduction to Psychology 1st Canadian Edition, https://opentextbc.ca/introductiontopsychology/chapter/4-2-seeing/ (accessed Apr. 14, 2024).

[5] “Diabetic retinopathy,” National Eye Institute, https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy (accessed Apr. 14, 2024).

[6] S. Aumann, “Optical Coherence Tomography (OCT): Principle and Technical Realization,” High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics [Internet]., https://www.ncbi.nlm.nih.gov/books/NBK554044/ (accessed Apr. 14, 2024).

[7] M. Karmel, “Retinal imaging: Choosing the right method,” American Academy of Ophthalmology, https://www.aao.org/eyenet/article/retinal-imaging-choosing-right-method (accessed Apr. 14, 2024).

[8] “Diabetic retinopathy,” Johns Hopkins Medicine, https://www.hopkinsmedicine.org/health/conditions-and-diseases/diabetes/diabetic-retinopathy (accessed Apr. 14, 2024).

[9] “Supervised vs. Unsupervised Learning: What's the difference?,” Google, https://cloud.google.com/discover/supervised-vs-unsupervised-learning (accessed Apr. 15, 2024).

[10] C. Doersch, Tutorial on Variational Autoencoders (2021), arXiv

[11] D. Fu, “Advances in understanding, improving, and applying contrastive learning,” Hazy Research, https://hazyresearch.stanford.edu/blog/2022-04-19-contrastive-1 (accessed Apr. 15, 2024).

[12] [1] L. Weng, “Contrastive Representation Learning,” Lil’Log, https://lilianweng.github.io/posts/2021-05-31-contrastive/ (accessed Apr. 15, 2024).

[13] J. Kukačka, A. Zenz, M. Kollovieh, D. Jüstel, and V. Ntziachristos, Self-Supervised Learning from Unlabeled Fundus Photographs Improves Segmentation of the Retina. 2021.

[14] N. Dimitriou, O. Arandjelović, and P. D. Caie, “Deep Learning for Whole Slide Image Analysis: An overview,” Frontiers in medicine, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6882930/ (accessed Apr. 15, 2024).\

[15] O. Ciga et al., “Overcoming the limitations of patch-based learning to detect cancer in whole slide images,” Nature News, https://www.nature.com/articles/s41598-021-88494-z (accessed Apr. 15, 2024).