CAP 5516
Medical Image Computing
(Spring 2025)

Dr. Chen Chen
Associate Professor
Center for Research in Computer Vision (CRCV)
University of Central Florida
Office: HEC 221
Email: chen.chen@crcv.ucf.edu
Web: https://www.crcv.ucf.edu/chenchen/

UCF CENTER FOR RESEARCH IN COMPUTER VISION

Lecture 5: Introduction to Medical Image Computing (3)

Image credit: https://www.syberscribe.com.au/blog/the-ins-and-outs-of-medical-imaging/

UCF CENTER FOR RESEARCH IN COMPUTER VISION

Pathology

Definition & Role:

- *Pathology:* The study of disease at the tissue, cellular, and molecular levels.
- *Importance:* Central to diagnosis, guiding treatment decisions, and understanding disease mechanisms.

Subfields of Pathology:

- Histopathology, cytopathology, molecular pathology, etc.

Pathology

Subfields of Pathology:

1. **Histopathology**

- Focus: Examination of *tissue sections* (e.g., from biopsies or surgical resections) under a microscope.
- Relevance: Histopathology is central to cancer diagnosis, among other conditions.

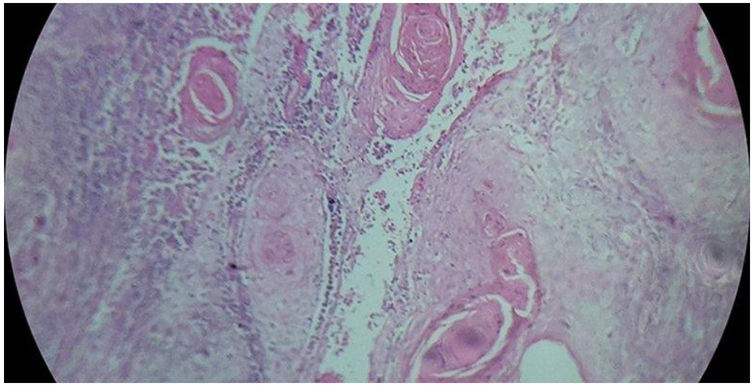


Image source: https://www.rcpath.org/discover-pathology/news/fact-sheets/histopathology.html

Pathology

Subfields of Pathology:

2. **Cytopathology**

- Focus: Study of *individual cells or small clusters of cells* (e.g., Pap smears, fine needle aspirations).
- Relevance: Useful for screening and diagnosing certain cancers (cervical, thyroid) and infections.

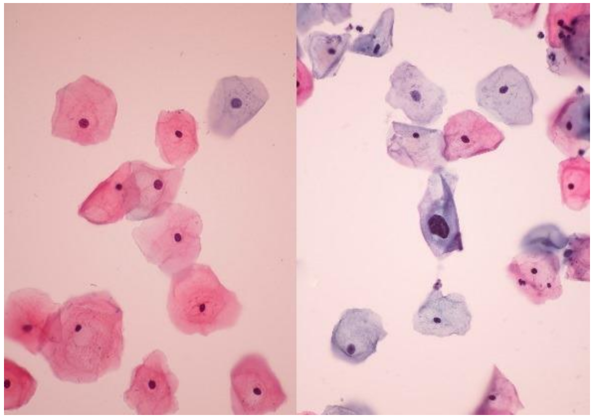


Image source: https://www.news-medical.net/life-sciences/What-is-cytopathology.aspx

Pathology

Subfields of Pathology:

3. **Molecular Pathology**

Molecular pathology is the study of diseases at the molecular level, using laboratory techniques to analyze the molecules found in tissues, organs, and bodily fluids.

- Focus: Analyzing *genes, proteins, and other molecules* within tissues or cells.
- Relevance: Essential for personalized medicine, targeted therapies, and advanced disease classification.

Histopathology

- **Key Use Case for WSI**: Whole Slide Imaging predominantly digitizes *tissue slides*, which is the foundation of histopathology.

- **Widespread Clinical Practice**: Most traditional pathology workflows revolve around tissue-based diagnoses (e.g., tumor grading, margin assessments).

- **Largest Potential Impact**: Since tissue slides are physically large and numerous in clinical labs, digitizing them offers massive benefits in storage, collaboration, and AI-driven analysis.

Traditional Microscopy in Pathology

Standard Workflow:

- Tissue biopsy or resection.
- Fixation, paraffin embedding, sectioning, staining (e.g., H&E).
- Examination under a light microscope.

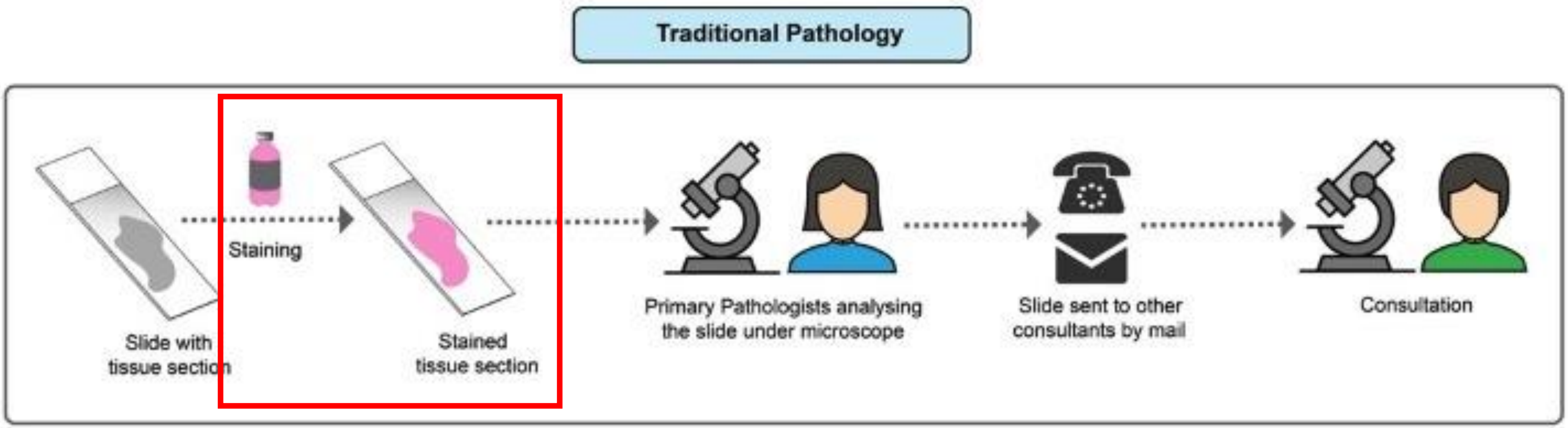


Image source: https://www.precisiononcology.ie/newsevents/blogs/items/text,502846,en.html

H&E Staining – The Gold Standard in Histopathology

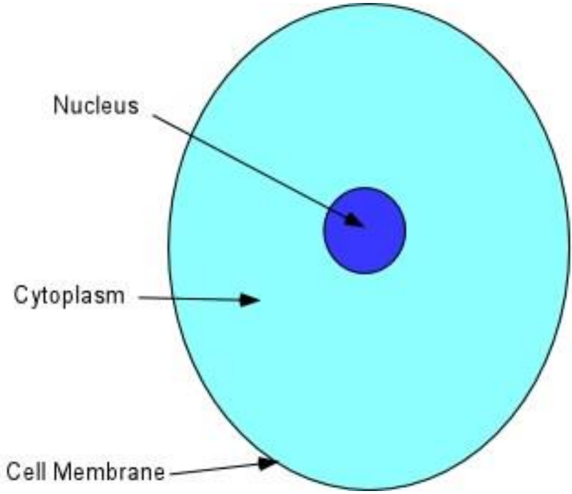
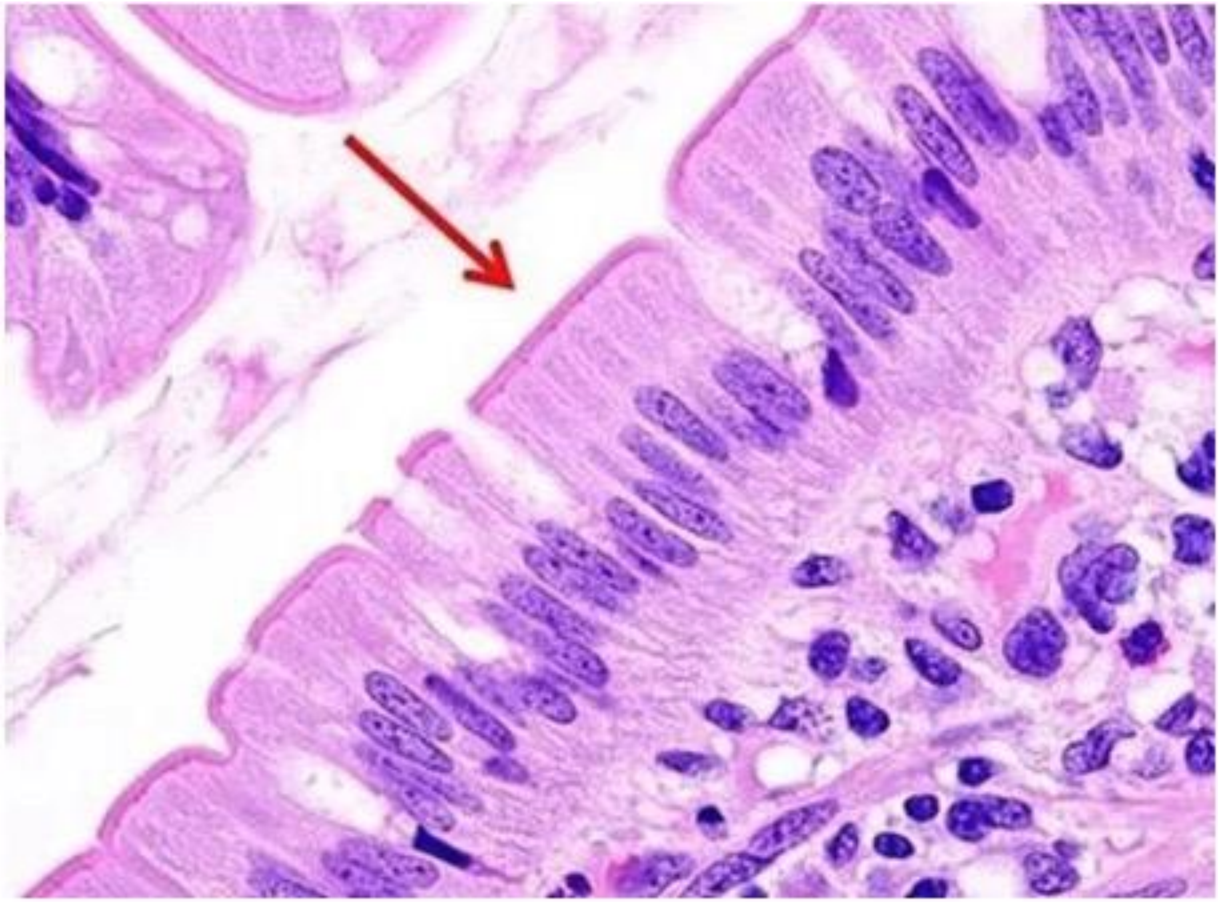


Image source: https://breast-cancer.ca/cytoplagarsm/

What is H&E Staining?

- **Hematoxylin** and **Eosin** are two dyes used in combination to stain tissue sections.
  - **Hematoxylin** (a basic dye) binds to **acidic components** (e.g., DNA in the cell nucleus), staining them **blue-purple**.
  - **Eosin** (an acidic dye) binds to **basic components** (e.g., proteins in the cytoplasm), staining them **pink**.
- **Enhanced Visual Contrast**
  - By assigning **distinct colors** to nuclei and cytoplasm, we can easily differentiate **cell boundaries** and **tissue architecture**.
  - This contrast is critical for identifying **morphological details** such as nuclear shape, size, and chromatin pattern.

H&E Staining – The Gold Standard in Histopathology



Colon H&E

Image source: https://www.leicabiosystems.com/us/knowledge-pathway/he-staining-overview-a-guide-to-best-practices/

H&E Staining – The Gold Standard in Histopathology

- **Universal Standard**: Most widely used staining method in pathology labs worldwide.
- **Cellular Detail**: Highlights nuclei (purple/blue) and cytoplasm (pink), providing clear contrast.
- **Broad Diagnostic Utility**: Suitable for a wide range of tissues and pathological conditions.
- **Routine & Cost-Effective**: Quick, reliable, and relatively inexpensive.

Traditional Microscopy in Pathology

Strengths:

- Established "gold standard" for diagnosis.
- Pathologists have extensive experience and comfort with glass slides.

Limitations:

- *Physical Slide Management:* Storage, risk of damage.
- *Limited Field of View:* Must manually move the slide under the microscope.
- *Collaboration Constraints:* Sending slides physically to consult with colleagues.
- *No Digital Data:* Hard to share or apply computational tools unless you photograph or scan manually.

Need for Digital Pathology

Modern healthcare demands:

- **Rapid Sharing** of diagnostic information.
- **Scalable Storage** for large numbers of cases.
- **Objective Analysis** (e.g., AI or quantitative analysis).

We require a more **robust, scalable, and analyzable** approach.

Digital Pathology

The Age of Digital Pathology

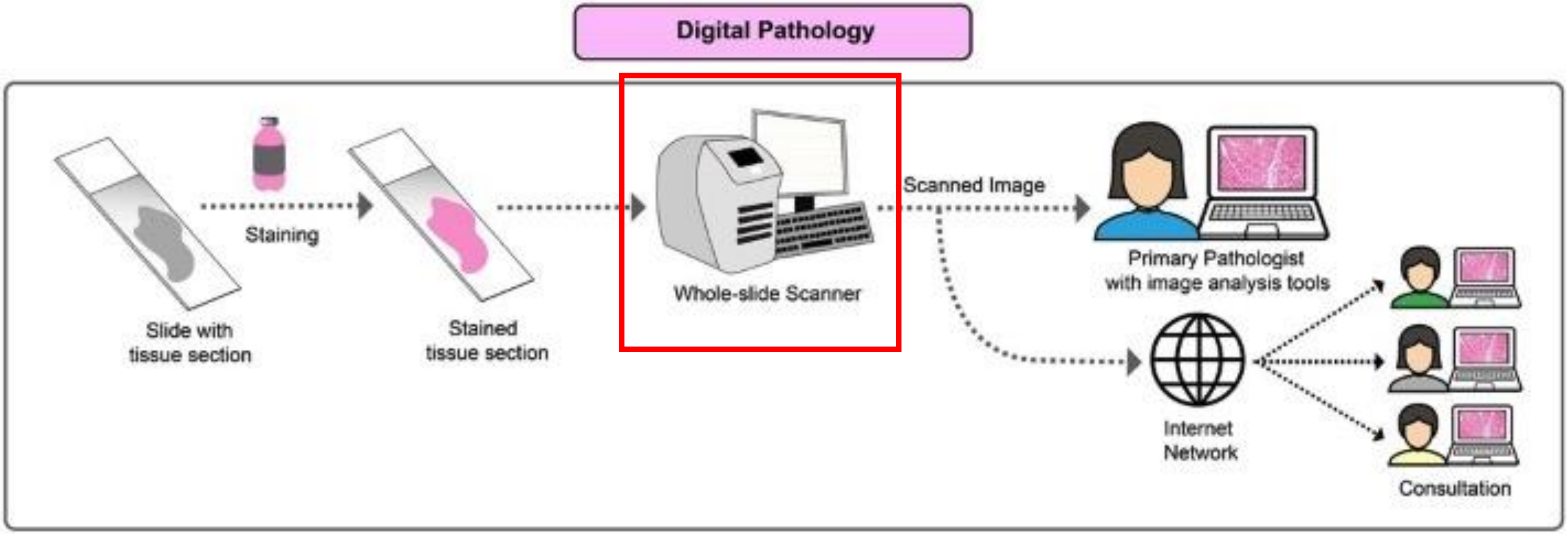


Image source: https://www.precisiononcology.ie/newsevents/blogs/items/text,502846,en.html

What is Whole Slide Imaging (WSI)?

Definition and Concept

- **Whole Slide Imaging (WSI)**: A process that digitizes an entire histology slide at high magnification, creating a virtual "digital slide."
- **Virtual Microscopy**: Rather than looking through a physical microscope, pathologists and researchers view these slides on a computer using specialized software.

Whole Slide Imaging (WSI)

How It Works (Basic Steps)

1. **Slide Preparation**: Tissue is fixed, stained (e.g., H&E), and coverslipped—just like with traditional microscopy.

2. **Scanning**:

- A high-speed scanner captures **tiled** images across the slide.
- Each tile overlaps slightly to ensure no gaps.
- **Optional**: Multiple focal planes (Z-stacking) if the tissue thickness or focus level is critical.

3. **Stitching**: Software combines these tiles into a single, large, high-resolution image.

4. **Digital File Creation**: The final image is saved in a specialized format (e.g., .svs, .ndpi, .ome.tiff).

Whole Slide Imaging (WSI)

How It Works (Basic Steps)

1. **Slide Preparation**: Tissue is fixed, stained (e.g., H&E), and coverslipped—just like with traditional microscopy.

2. **Scanning**:

- A high-speed scanner captures **tiled** images across the slide.
- Each tile overlaps slightly to ensure no gaps.
- **Optional**: Multiple focal planes (Z-stacking) if the tissue thickness or focus level is critical.

3. **Stitching**: Software combines these tiles into a single, large, high-resolution image.

4. **Digital File Creation**: The final image is saved in a specialized format (e.g., .svs, .ndpi, .ome.tiff).

Whole Slide Imaging (WSI)

Tiling

- **Why Tiling?**
    - **High Resolution Requirement:** Scanners need to capture the entire slide at 20x, 40x, or even higher magnification. A single snapshot cannot cover the entire slide due to the limited field of view of the microscope lens.
    - **Efficient Scanning:** By capturing many small, high-magnification "tiles," the scanner can systematically move across the slide (like a grid) without missing any regions.
- **How Tiling Works**
    - **Motorized Stage Movement:** The slide is placed on a **motorized stage** that moves incrementally in the **x** and **y** directions.
    - **Capturing Tiles:** For each position, the scanner takes a **high-resolution image** of a small region—this is one "tile."
    - **Overlap for Accuracy:** Tiles typically overlap by a small margin (e.g., 5–10%) to ensure complete coverage and aid in alignment.

Whole Slide Imaging (WSI)

One Plane vs. Multiple Planes:

- In a standard WSI, the scanner captures the slide at a single focal depth, producing one high-resolution image.
- When a slide scanner captures **multiple focal planes**, it effectively creates a **“stack” of images** for the same tissue region at different depths (z-positions). With **Z-stacking**, each focal depth is like an additional image layer.
- **Data Multiplication:** If you have n focal planes, you’re essentially acquiring n high-resolution images for each tile on the slide.

Multiple Planes

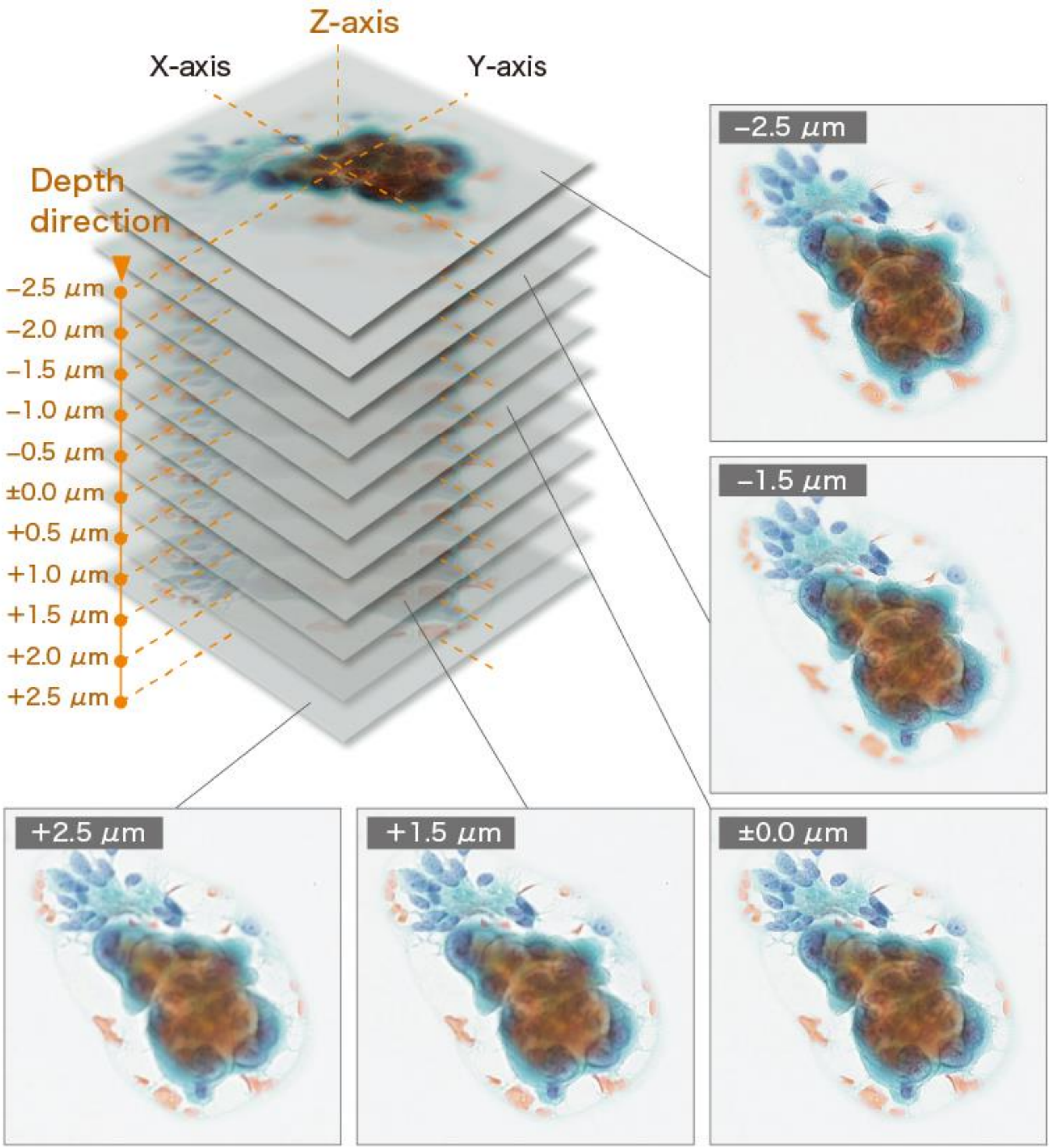


Image source: https://nanozoomer.hamamatsu.com/jp/en/why_nanozoomer/scan.html

Whole Slide Imaging (WSI)

- The **majority** of whole slide images (WSIs) are stored in a **pyramidal** (multi-resolution) format.

- **Pyramidal File Structure:** A **pyramidal file** (sometimes called a "multi-resolution pyramid") stores **multiple versions** (resolutions) of the same image within a single file.

- **How It's Created:** After the highest resolution image is captured and stitched, the software **downsamples** it to create progressively lower-resolution images, each level being a fraction of the previous size (e.g., 1/2, 1/4, 1/8).

Whole Slide Imaging (WSI)

Pyramid Layers (Zoom Levels)

- **Highest Level (Full Resolution):** Each pixel captured by the scanner. Extremely detailed, but can be gigabytes in size.

- **Intermediate Levels:** Several "middle" layers.

- **Lowest Level (Thumbnail):** A very small version for quick preview.

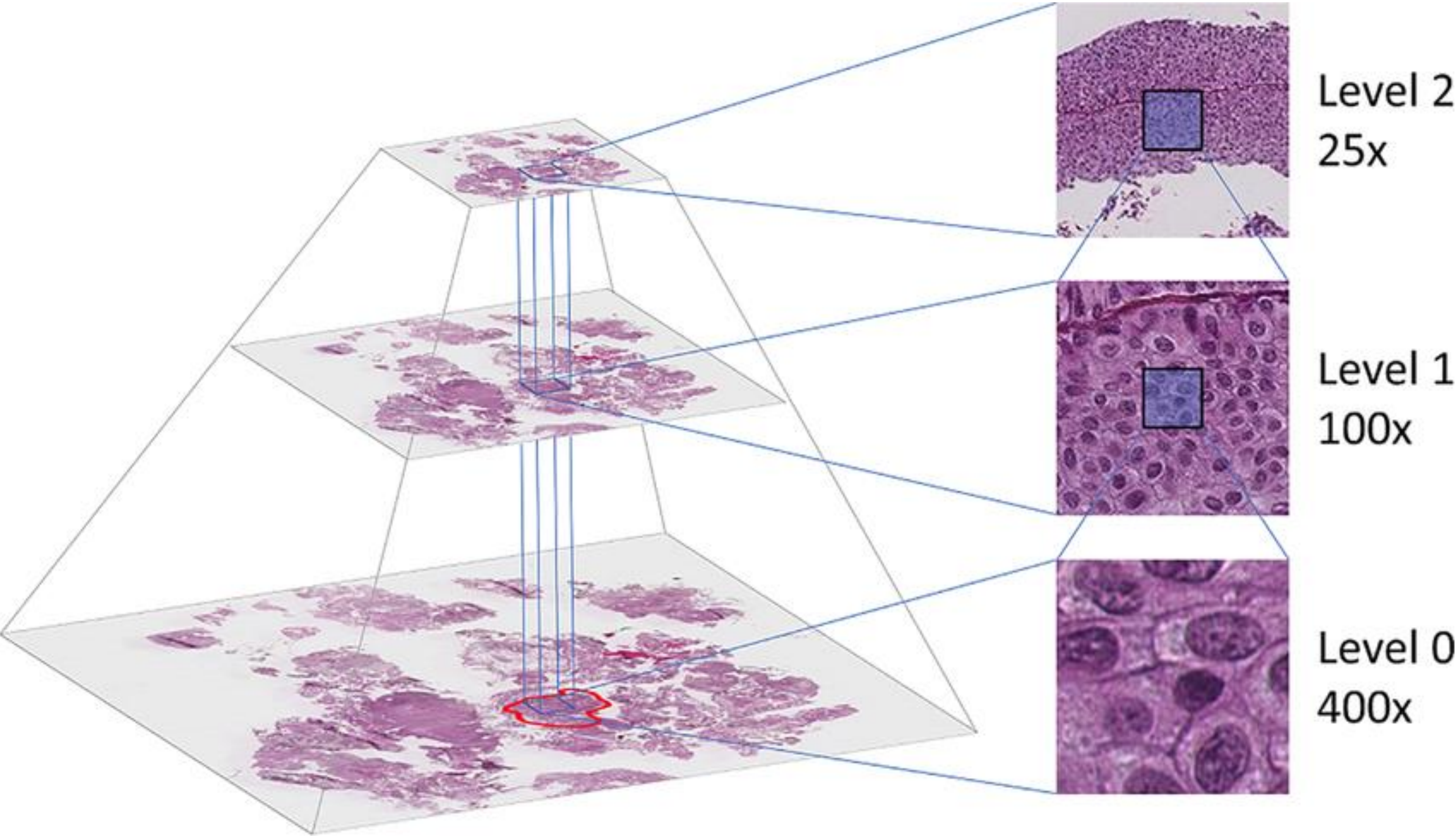


Image source: Wetteland, Rune, et al. "A multiscale approach for whole-slide image segmentation of five tissue classes in urothelial carcinoma slides." *Technology in Cancer Research & Treatment* 19 (2020): 1533033820946787.

Whole Slide Imaging (WSI)

Common WSI File Formats

- **Proprietary Formats**
  - **.svs (Aperio)**, **.ndpi (Hamamatsu)**, **.vms / .vmu (Ventana)**, **.mrxs (3DHISTECH)**
  - Often optimized for each vendor's scanner, which can complicate cross-platform compatibility.
- **Open Formats**
  - **OME-TIFF (Open Microscopy Environment)**
  - Aims to standardize metadata and improve interoperability.
  - Supports multi-channel (e.g., fluorescence) and multi-plane imaging.
- **DICOM** for Digital Pathology
  - Digital Imaging and Communications in Medicine (DICOM) standard adapted to pathology.
  - Still gaining traction but promises improved interoperability and integration with hospital PACS systems.
  - https://dicom.nema.org/dicom/dicomwsi/

Whole Slide Imaging (WSI)

Software for Viewing WSI

1. **QuPath (Open Source)**

• **Platform:** Cross-platform (Windows, Mac, Linux).

• **Features:** Zooming, panning, annotation, basic image analysis, and scripting.

• **Website:** https://qupath.github.io/

• **Great For:** Teaching basic digital pathology concepts, creating annotations, and even simple AI demos.

2. **OpenSlide (Library & Viewers)**

• **OpenSlide** is a C library that allows you to read many proprietary WSI formats.

• **OpenSlide Python** or **OpenSlide Java** can be used to build custom viewers or scripts.

• They also have a simple viewer tool called openslide-python's *DeepZoom* demo.

Website: https://openslide.org/

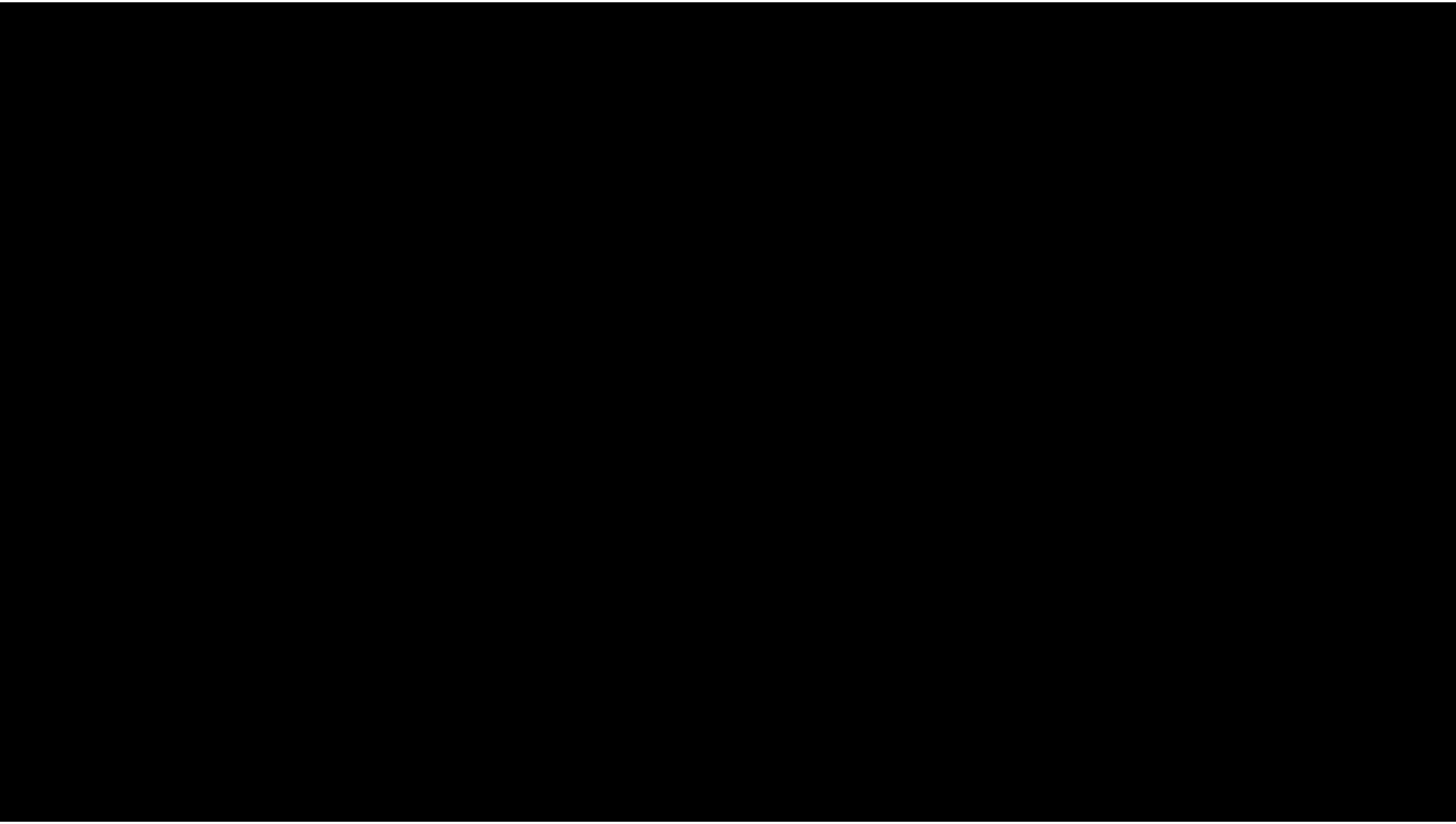
3. **PathPresenter (Online Viewer)**

- Demo

https://openslide.org/demo/

WSI Scanner

Image source: https://evidentscientific.com/en/news/slideview-dx-quality-images-fast-efficient-digital-pathology#:~:text=WALTHAM%2C%20Mass.%2C%20(June,make%20diagnoses*%20quickly%20and%20efficiently.



Video source: https://www.youtube.com/watch?v=6X5K7w5jeFU

Key AI Applications in Whole Slide Imaging (WSI)

Automated Tumor Detection and Grading

- **Detection:** AI algorithms identify and locate tumor regions within WSIs.
- **Grading:** Assess tumor aggressiveness (e.g., Gleason score in prostate cancer).

Benefits:

- **Increased Accuracy:** Reduces human error and inter-observer variability.
- **Efficiency:** Speeds up diagnostic workflow, allowing pathologists to focus on complex cases.

Key AI Applications in Whole Slide Imaging (WSI)

Cell Segmentation and Classification

- **Segmentation:** AI delineates individual cells or nuclei within tissue samples.
- **Classification:** Each segmented cell is categorized (e.g., cancerous vs. non-cancerous).

Benefits:

- **Detailed Analysis:** Facilitates quantitative studies of cellular populations.
- **Scalability:** Efficiently handles large datasets essential for high-resolution WSIs.

Questions?

Ophthalmology and Optometry

Image source: https://eduadvisor.my/articles/optometrist-vs-ophthalmologist-whats-the-difference

Ophthalmology

Definition:

- A medical and surgical specialty focused on the diagnosis, treatment, and management of eye diseases and conditions.

Practitioners:

- Ophthalmologists are medical doctors (MDs or DOs) who perform eye surgeries, prescribe medications, and manage complex eye diseases.

Scope:

- Comprehensive eye care, including surgical interventions, advanced diagnostics, and treatment of serious ocular conditions.

Ophthalmology and Optometry

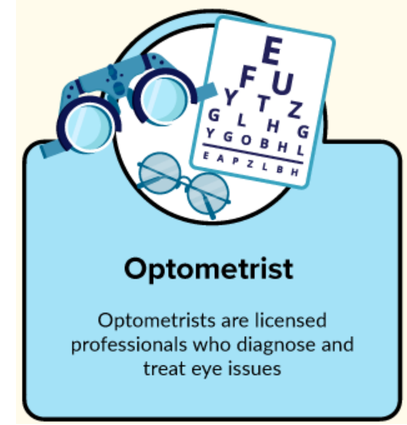


Image source: https://eduadvisor.my/articles/optometrist-vs-ophthalmologist-whats-the-difference

Optometry

Definition:

- A healthcare profession dedicated to vision care, including eye examinations, prescribing corrective lenses, and diagnosing certain eye conditions.

Practitioners:

- Optometrists (ODs) provide primary vision care, perform routine eye exams, and prescribe glasses and contact lenses.

Scope:

- Focus on vision correction, management of refractive errors, and detection of common eye diseases; may refer patients to ophthalmologists for advanced care.

Overview of Key Ophthalmic Imaging Modalities

- **Main Imaging Modalities:**

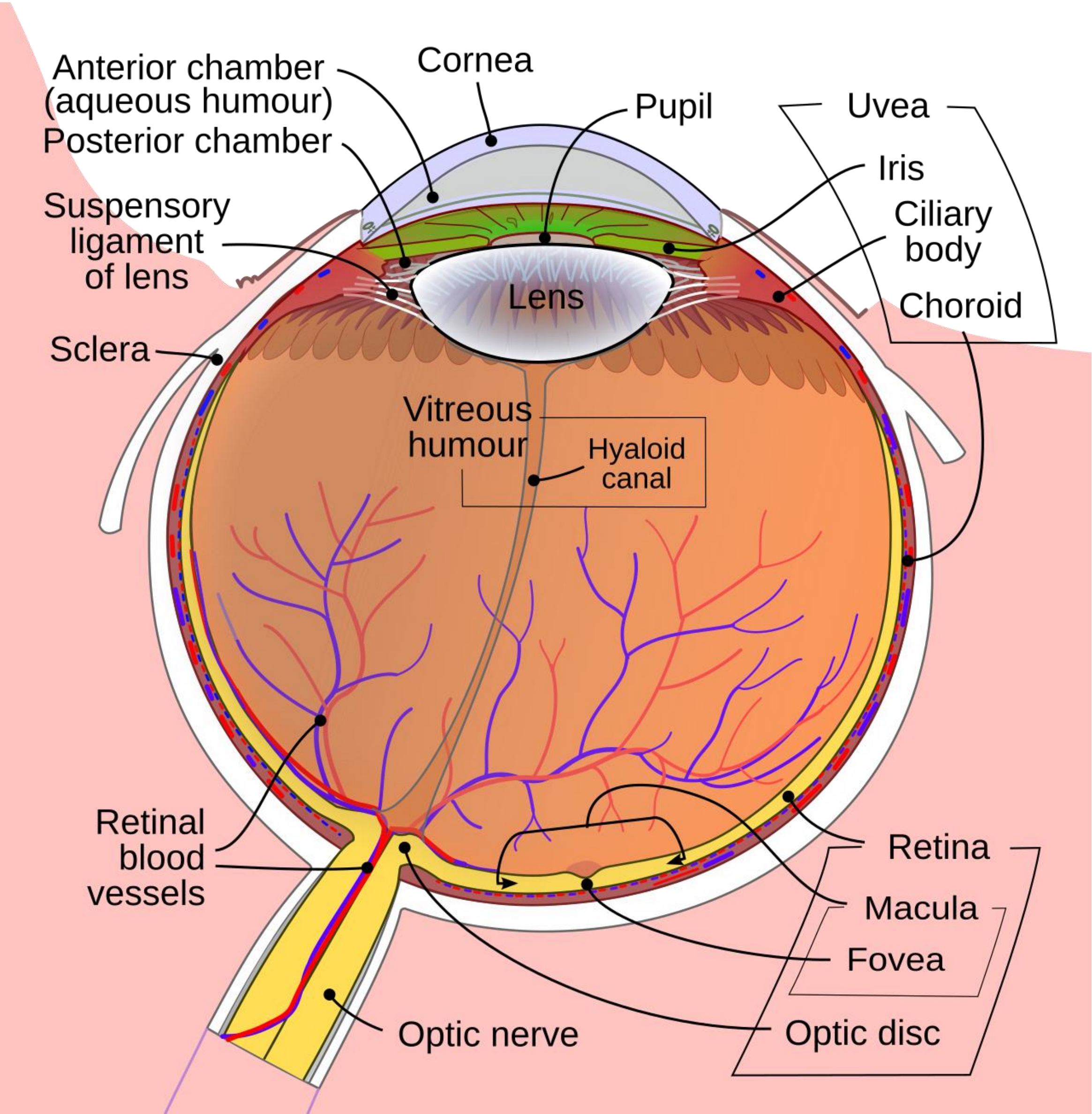
1. Fundus Photography
2. Fluorescein Angiography (FA)
3. Optical Coherence Tomography (OCT)

Overview of Key Ophthalmic Imaging Modalities

Purpose of These Modalities:

- Provide detailed visualization of the eye's internal structures.
- Assist in early detection and precise diagnosis of ocular diseases.
- Enable monitoring of disease progression and treatment efficacy.

Fundus



The fundus of the eye is the back surface of the eye, and is made up of the retina, optic disc, macula, fovea, and blood vessels.

Image source: https://en.wikipedia.org/wiki/Eye

Fundus

Why We Care About the Fundus

- **Central to Vision Health**
  - **Retina:** Converts light into neural signals, essential for vision.
  - **Macula:** Responsible for sharp, central vision and color perception.
  - **Optic Disc:** Transmits visual information from the eye to the brain.

- **Indicator of Systemic Health**
  - **Diabetic Retinopathy:** Detects early signs of diabetes-related eye damage.
  - **Hypertensive Retinopathy:** Monitors blood pressure effects on retinal vessels.
  - **Age-Related Macular Degeneration (AMD):** Identifies progressive loss of central vision.

Fundus

Why We Care About the Fundus

- **Early Detection of Ocular Diseases**
  - **Glaucoma:** Assesses optic nerve damage to prevent irreversible vision loss.
  - **Retinal Vein Occlusion:** Identifies blockages in retinal blood vessels.
  - **Infections and Inflammatory Conditions:** Spot signs of uveitis or other inflammatory diseases.
- **Neurological Insights**
  - **Papilledema:** Detects increased intracranial pressure through optic disc swelling.
  - **Multiple Sclerosis:** Observes optic neuritis and other nerve-related changes.
- **Cardiovascular Health**
  - **Atherosclerosis:** Visualizes changes in retinal arteries as markers for heart disease.
  - **Cholesterol Deposits:** Identifies xanthomas or other lipid-related deposits in the eye.

Fundus Photography

A photographic technique that captures detailed images of the **fundus**, including the **retina**, **optic disc**, and **macula**.

How It Works:

- Utilizes a specialized camera with a bright light source and lens to photograph the interior surface of the eye.
- Patients may need to dilate their pupils for clearer images.

Fundus Camera

Image source: https://marco.com/product/afc-330-automated-fundus-camera/

Fundus Photography

Fundus photographs of the right eye (left image) and left eye (right image)

Image source: https://en.wikipedia.org/wiki/Fundus_(eye)

Key Applications of Fundus Photography

Diabetic Retinopathy Detection

Diabetic retinopathy (DR) is an eye condition that occurs when high blood sugar levels damage blood vessels in the retina.

- **Purpose:** Identify and grade diabetic retinopathy (DR) to prevent vision loss.
- **Significance:** Early detection of DR can lead to timely treatment, reducing the risk of severe complications like blindness.

AI Applications:

- Automated grading of DR severity.
- Detection of microaneurysms, hemorrhages, and exudates.

Available Datasets:

- **Messidor:** https://www.adcis.net/en/third-party/messidor/
- **Messidor 2:** https://www.adcis.net/en/third-party/messidor2/
- **APTOS (Asia Pacific Tele-Ophthalmology Society) 2019 Blindness Detection:**

https://www.kaggle.com/competitions/aptos2019-blindness-detection

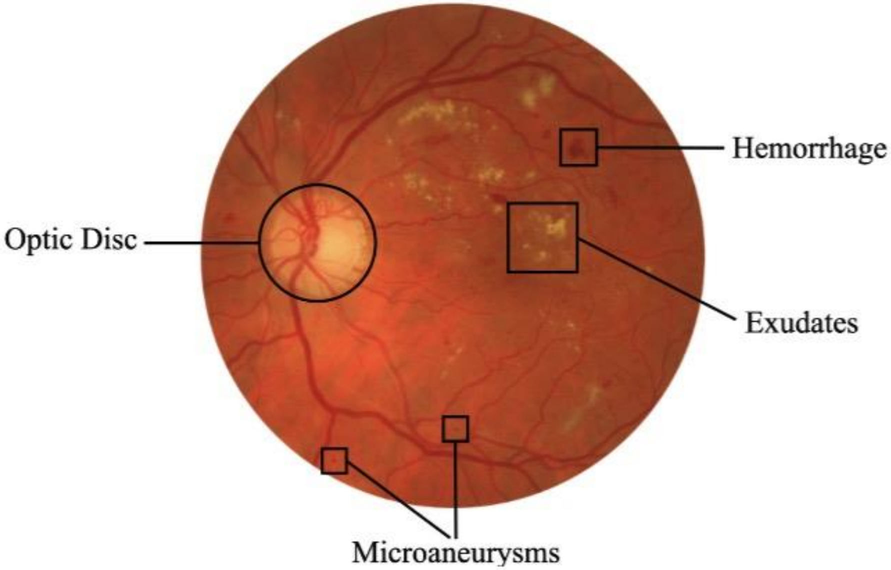


Figure 2: Fundus image with various lesions for DR classification.

Diabetic retinopathy (DR) is an eye condition that occurs when high blood sugar levels damage blood vessels in the retina.

Image source: Shakibania, H., Raoufi, S., Pourafkham, B., Khotanlou, H., & Mansoorizadeh, M. (2024). Dual branch deep learning network for detection and stage grading of diabetic retinopathy. Biomedical Signal Processing and Control, 93, 106168.

- **Microaneurysms:** Small bulges in blood vessels indicating leakage points.
- **Hemorrhages:** Bleeding within the retinal layers.
- **Exudates:** Deposits of lipids and proteins from leaking vessels.
- **Neovascularization:** Formation of new, fragile blood vessels.

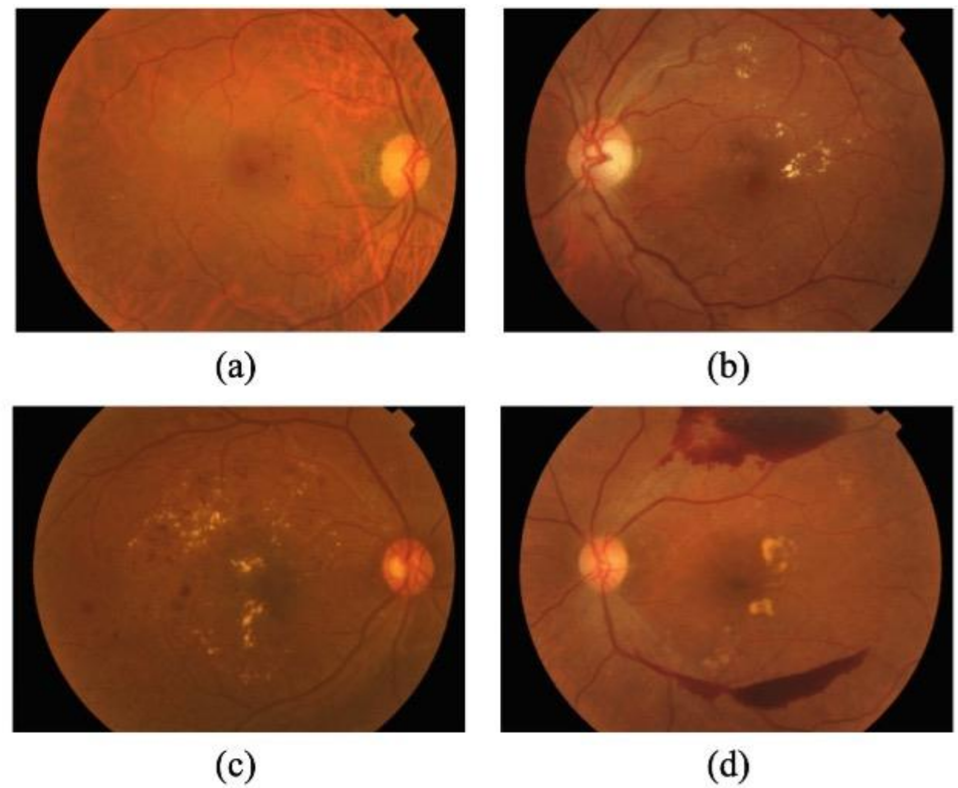


Figure 1: Four stages of DR: (a) Mild, (b) Moderate, (c) Severe, and (d) PDR obtained from APTOS 2019 dataset.

Key Applications of Fundus Photography

Glaucoma Assessment

Glaucoma is a chronic eye disease that damages the optic nerve, which can lead to vision loss and blindness

- Detect and monitor glaucoma by analyzing the optic nerve head and cup-to-disc ratio.

AI Applications:

- Automated segmentation of optic disc and cup.
- Classification of images as glaucomatous or healthy.

Available Datasets:

- **RIM-ONE (Retinal Images for Glaucoma Analysis): https://github.com/miag-ull/rim-one-dl**
- **Drishti-GS (Glaucoma Segmentation): https://cvit.iiit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php**

Fluorescein Angiography (FA): Visualizing Retinal Blood Flow

FA is an imaging technique that uses a fluorescent dye (fluorescein) to visualize the blood flow in the retina and choroid.

It can assist with evaluating the condition of the blood vessels in your eye and the circulation of your blood flow in the retina and choroid.

How It Works:

- **Procedure:**

1. **Injection:** Fluorescein dye is injected into a vein, usually in the arm.
2. **Imaging:** A specialized camera captures sequential images as the dye circulates through the retinal and choroidal blood vessels.
3. **Phases:** Images are taken at various stages to assess different phases of blood flow and leakage.

Fluorescein Angiography (FA)

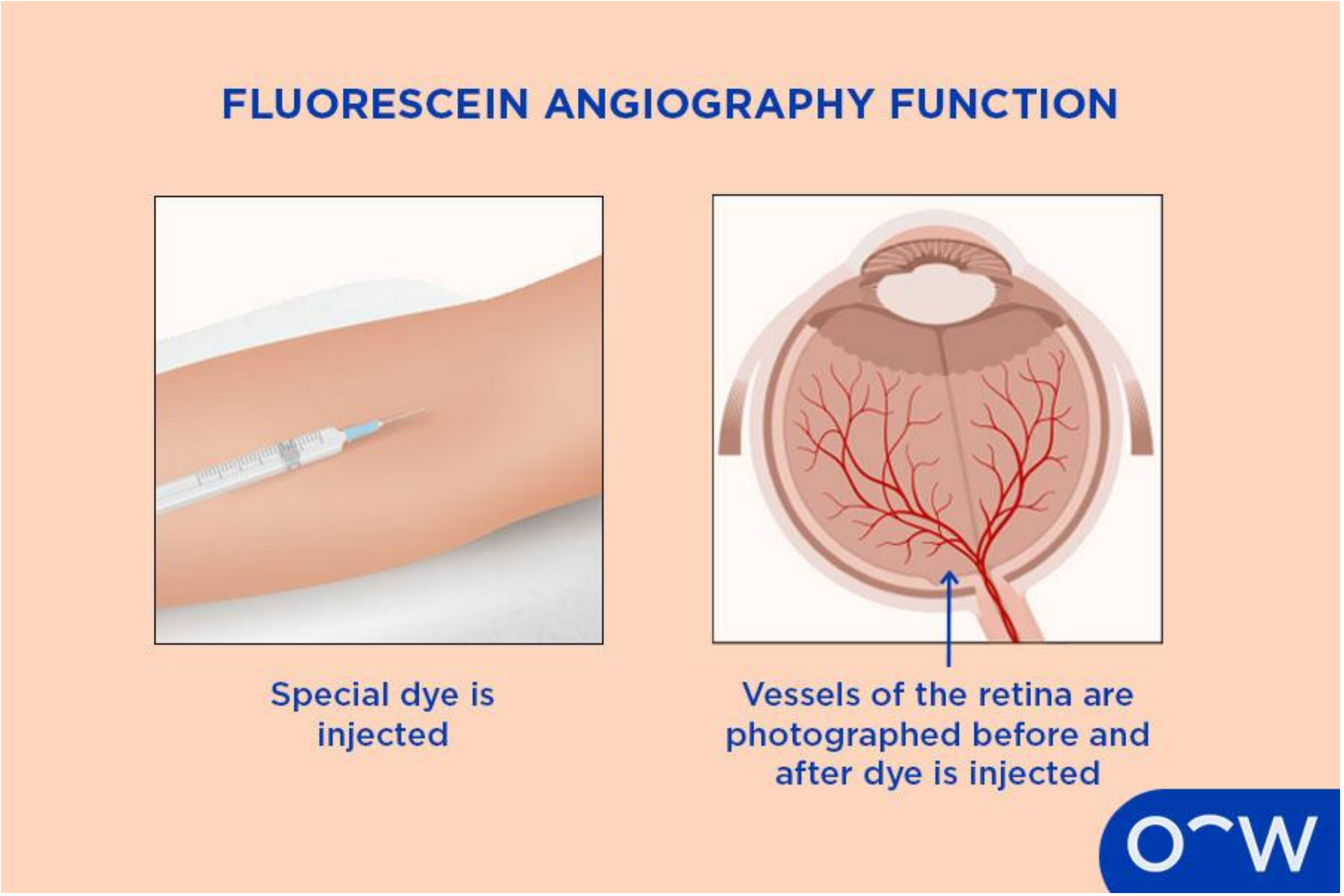
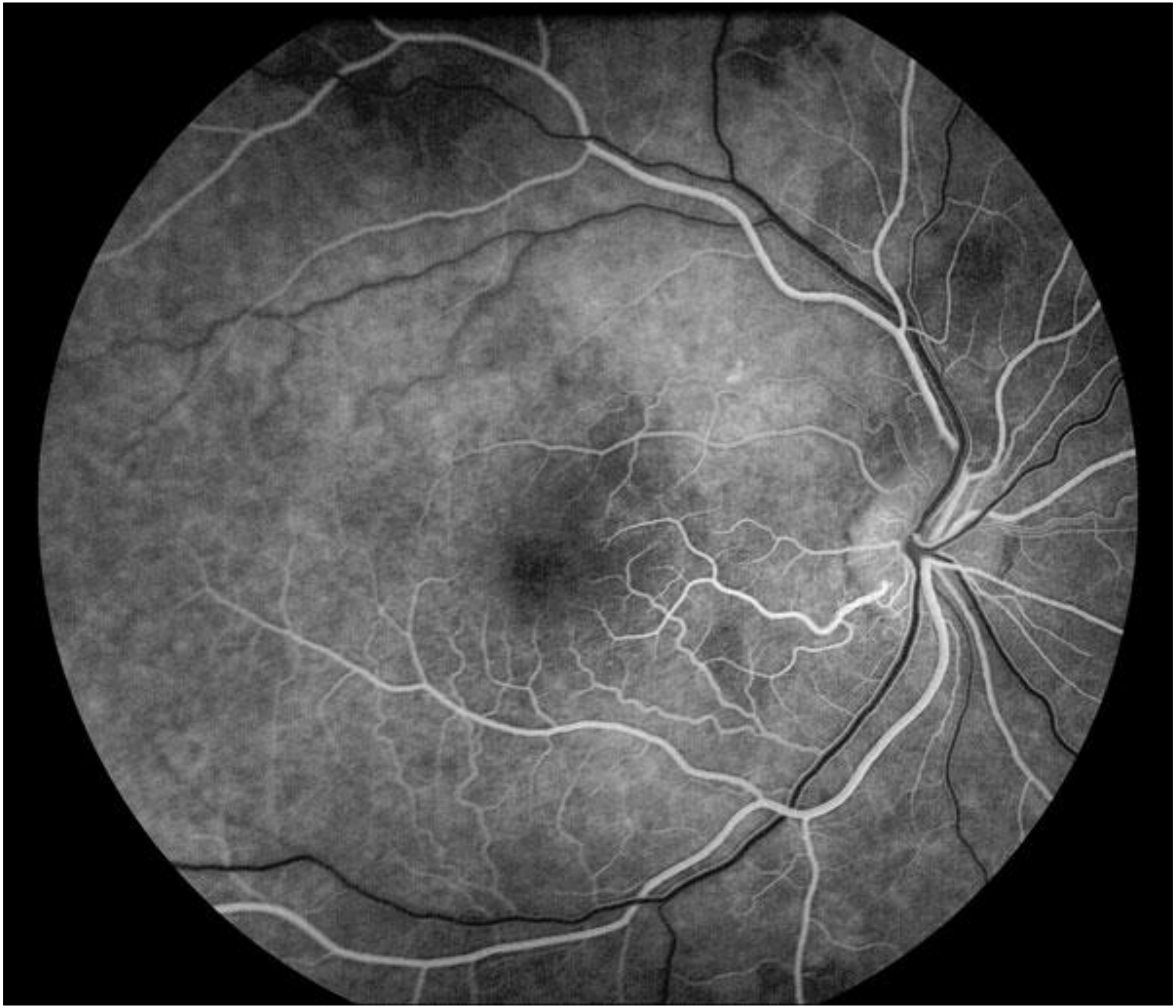


Image source: https://www.oscarwylee.com.au/glasses/eye/fluorescein-angiography

Fluorescein Angiography (FA)

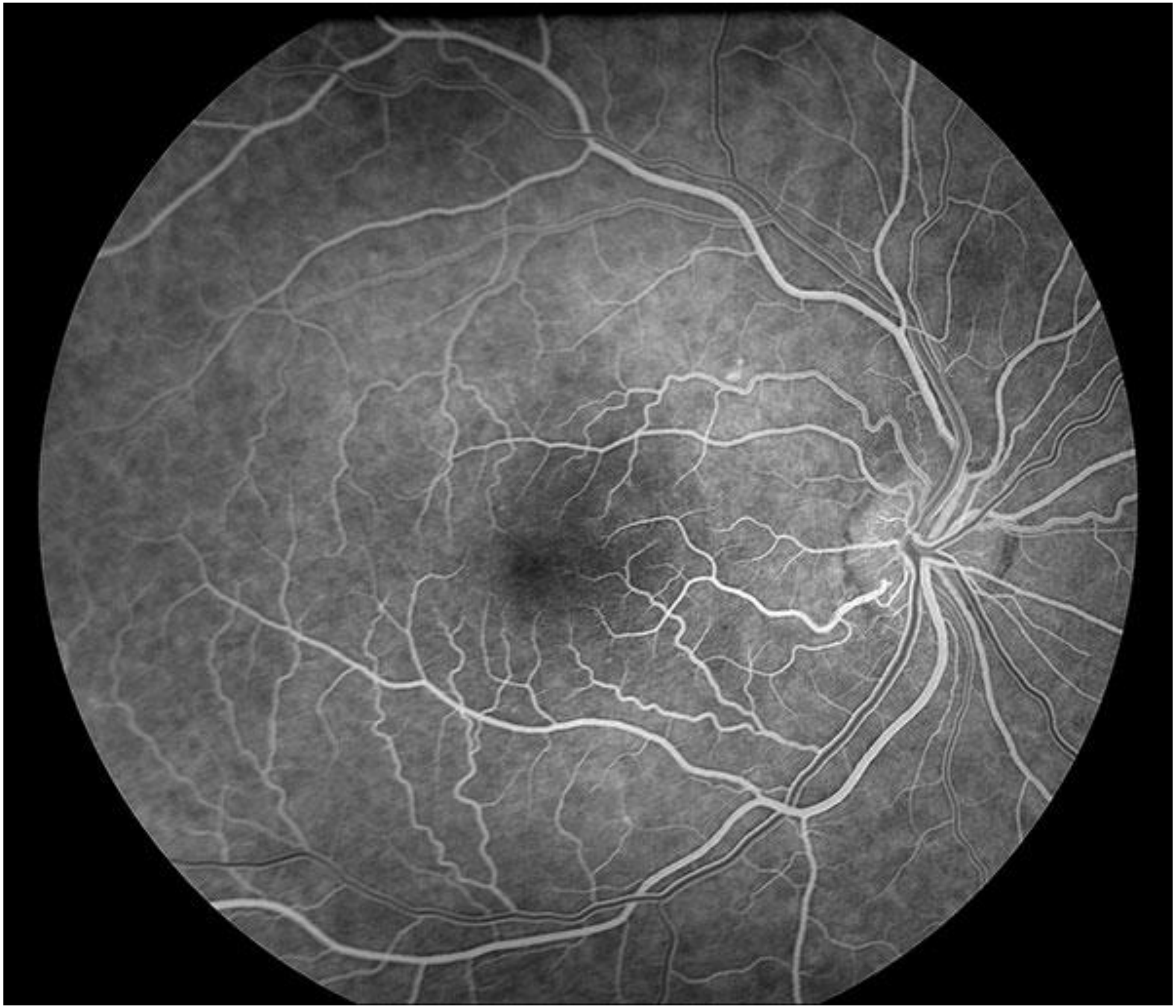
- The stages of fluorescein angiography **include the choroidal phase, arterial phase, arteriovenous, venous and recirculation**. The stages of fluorescein angiography are listed below.
- **Choroidal Phase:** The choroidal phase is where the dye takes under 20 seconds to reach the retina. The posterior ciliary arteries is the first region that the dye fills, therefore establishing a choroidal filling according to Ruia, S. and Tripathy, K. (2023).
- **Arterial Phase:** The arterial phase refers to when the dye moves to the arteries which can be as quick as 1 or 2 seconds after the choroidal phase.
- **Venous:** Venous involves three phases, early, mid and late depending on the filling of the fluorescein dye. The early phase exhibits a smooth flow, and then complete fulfilment of the dye in the veins in the mid phase, which is then followed by less dye in the arteries during the late phase.
- **Recirculation:** Recirculation is the occurrence after the venous phase, wherein, the dye begins reducing in the vessels and clearing out at around 10 minutes according to the National Library of Medicine.

Source: https://www.oscarwylee.com.au/glasses/eye/fluorescein-angiography



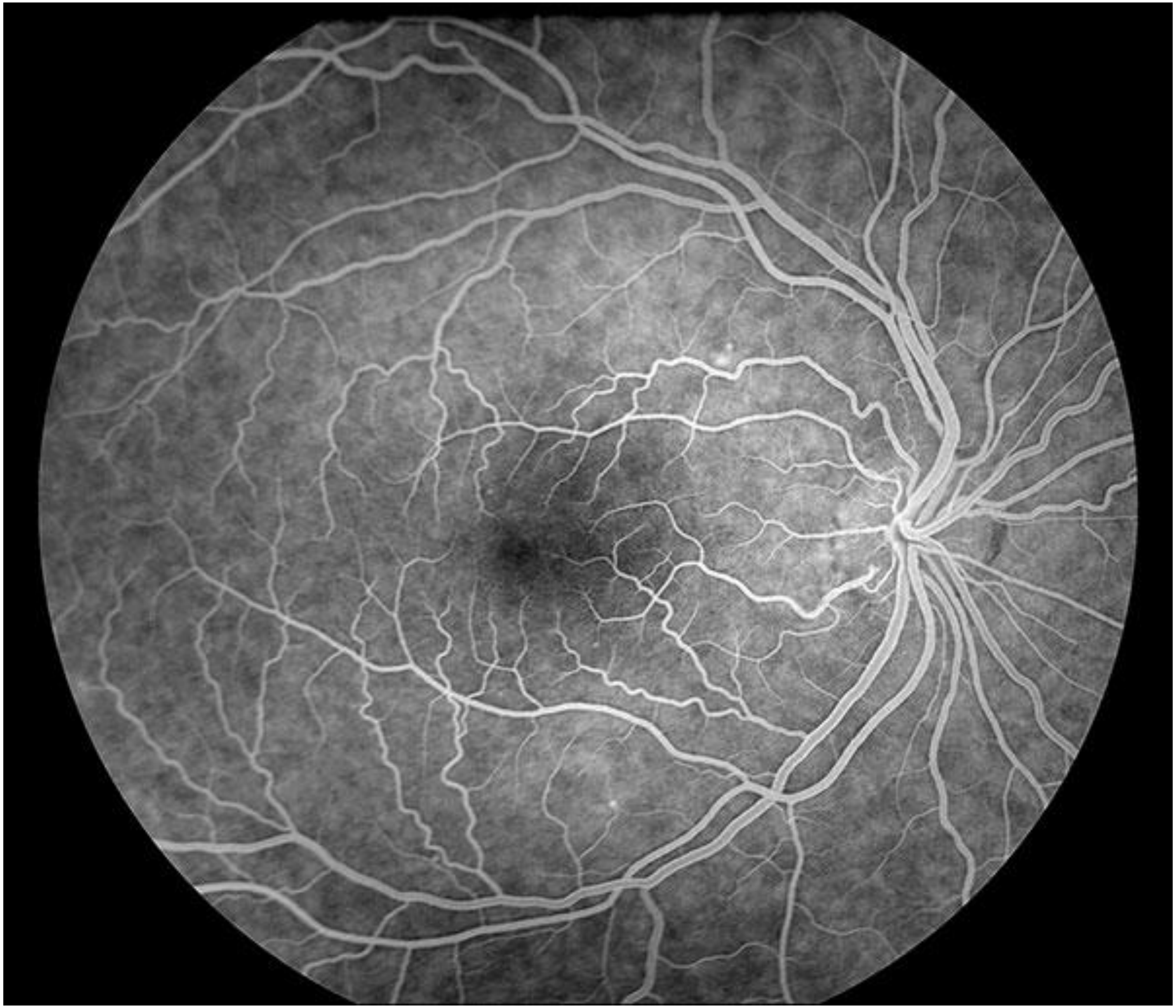
Fluorescein Angiogram. **Arterial phase** illustrating sodium fluorescein dye in the retinal arteries before filling the retinal vei*ns.*

Image source: https://medicine.uiowa.edu/eye/fluorescein-angiography



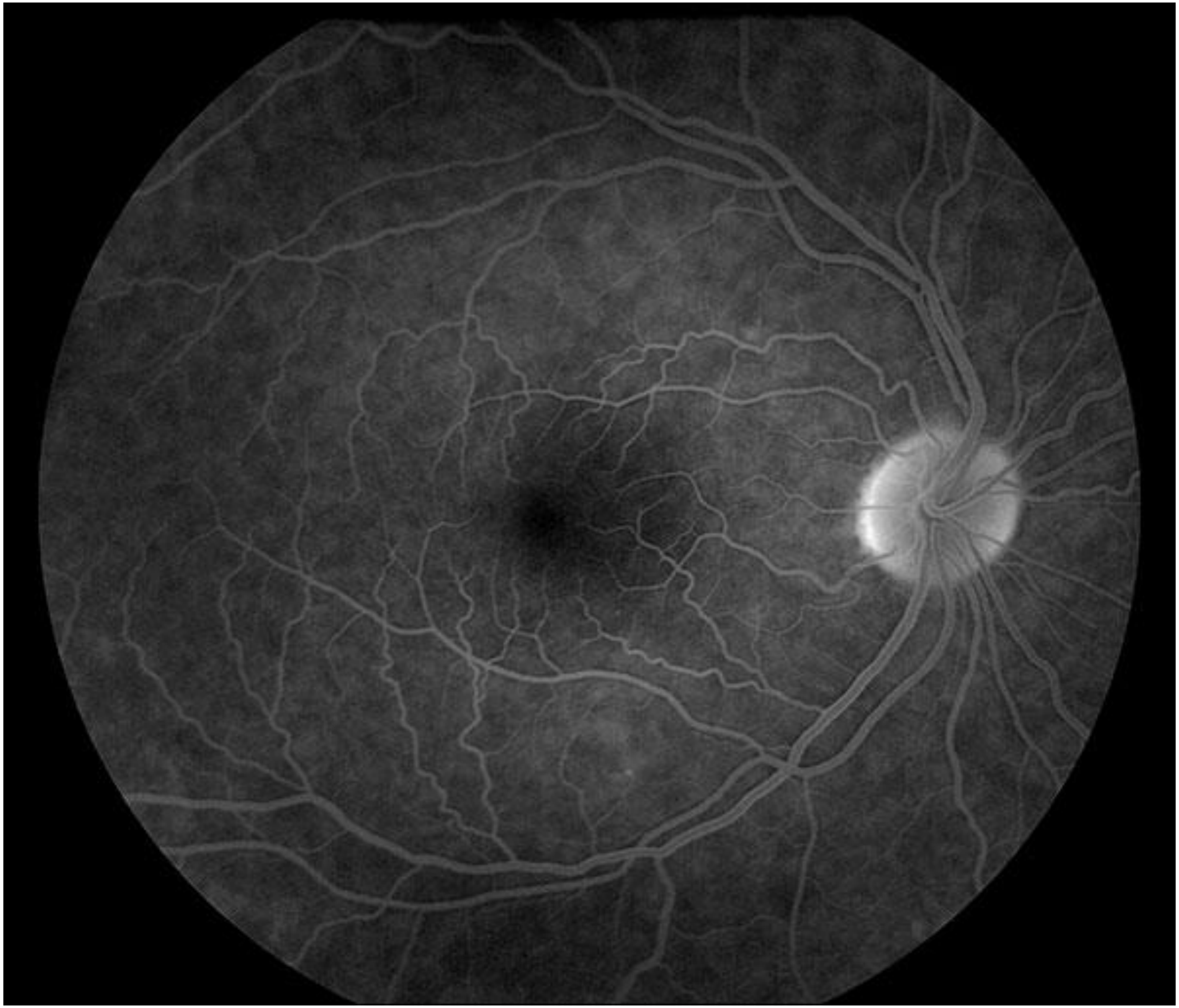
Early venous phase illustrating sodium fluorescein dye beginning to fill the retinal ve*ins.*

Image source: https://medicine.uiowa.edu/eye/fluorescein-angiography



Complete fill of retinal arteries and veins with sodium fluorescein dye.

Image source: https://medicine.uiowa.edu/eye/fluorescein-angiography



The **late phase** of the angiogram demonstrates **fading of the sodium fluorescein dye in the retinal vessels**.

Image source: https://medicine.uiowa.edu/eye/fluorescein-angiography

Applications of Fluorescein Angiography (FA)

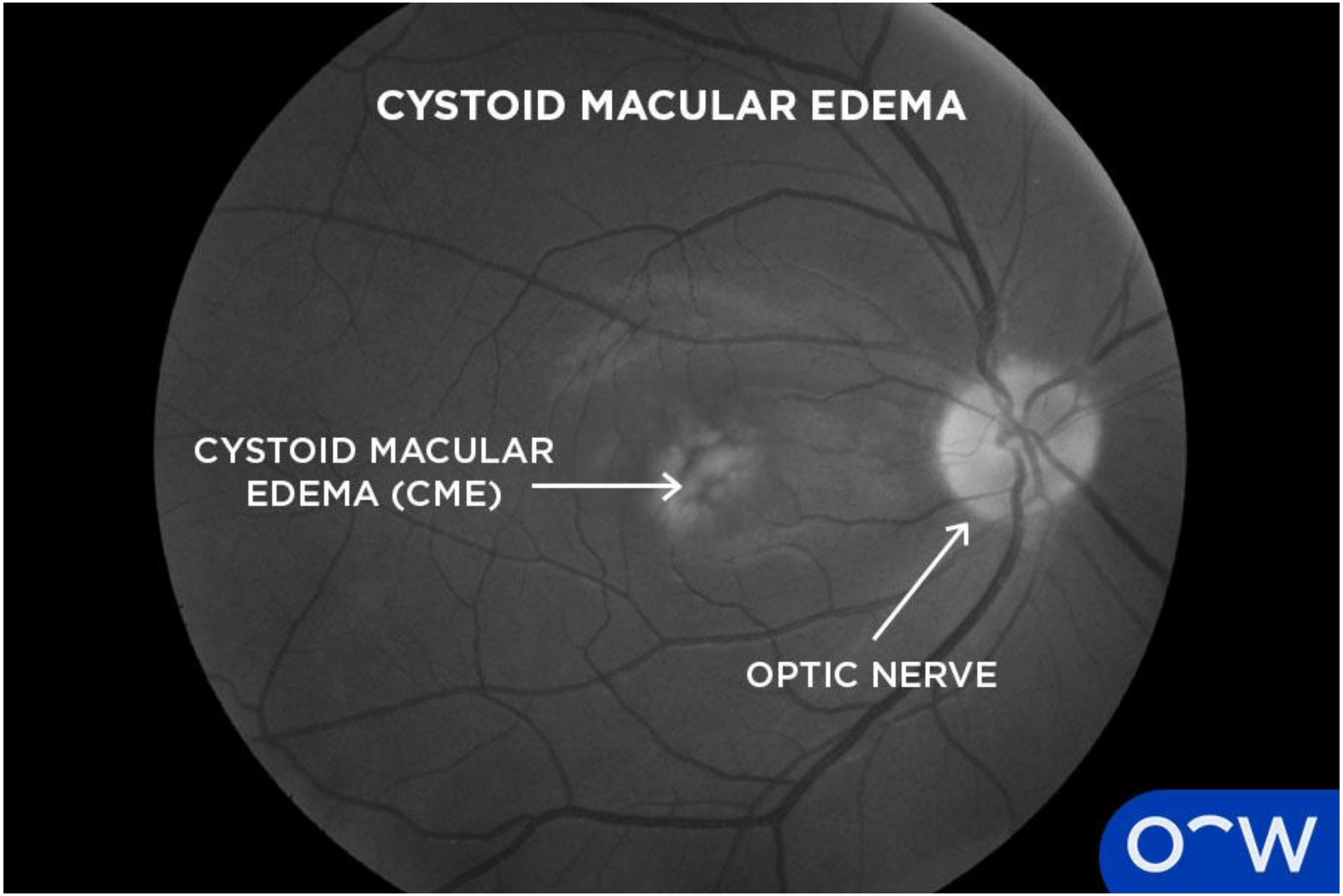
- Cystoid Macular Edema (CME) is a condition where **fluid accumulates in cyst-like spaces** within the **macula**, this leads to retinal swelling and distortion of vision.

Pathophysiology of CME:

- **Disruption of the Blood-Retinal Barrier:**
- Normally, the blood-retinal barrier prevents fluid leakage into retinal tissues.
- CME occurs when this barrier is compromised due to inflammation, ischemia, or vascular damage.
- **Vascular Leakage:**
- Leakage of fluid and proteins from retinal capillaries accumulates in the outer plexiform layer of the macula, forming cystic spaces.

How CME can be detected

- **Fluorescein Angiography (FA):**
  - Highlights **vascular leakage** and cystic spaces in the macula.
  - Shows characteristic **petaloid or honeycomb patterns** due to dye pooling.



Cystoid Macular Edema (CME), a condition characterized by the accumulation of fluid in the macula

Bright Regions in the Macula: The bright spots represent **fluorescein dye leakage**, suggesting damaged or dysfunctional blood-retinal barriers.

Image source: https://www.oscarwylee.com.au/glasses/eye/fluorescein-angiography

Optical Coherence Tomography (OCT)

- OCT is a **non-invasive imaging modality** that provides **high-resolution, cross-sectional images** of the retina and other ocular structures.

- OCT allows eye specialists to examine your eye in layers and to measure the depth of important structures.

Optical Coherence Tomography (OCT)

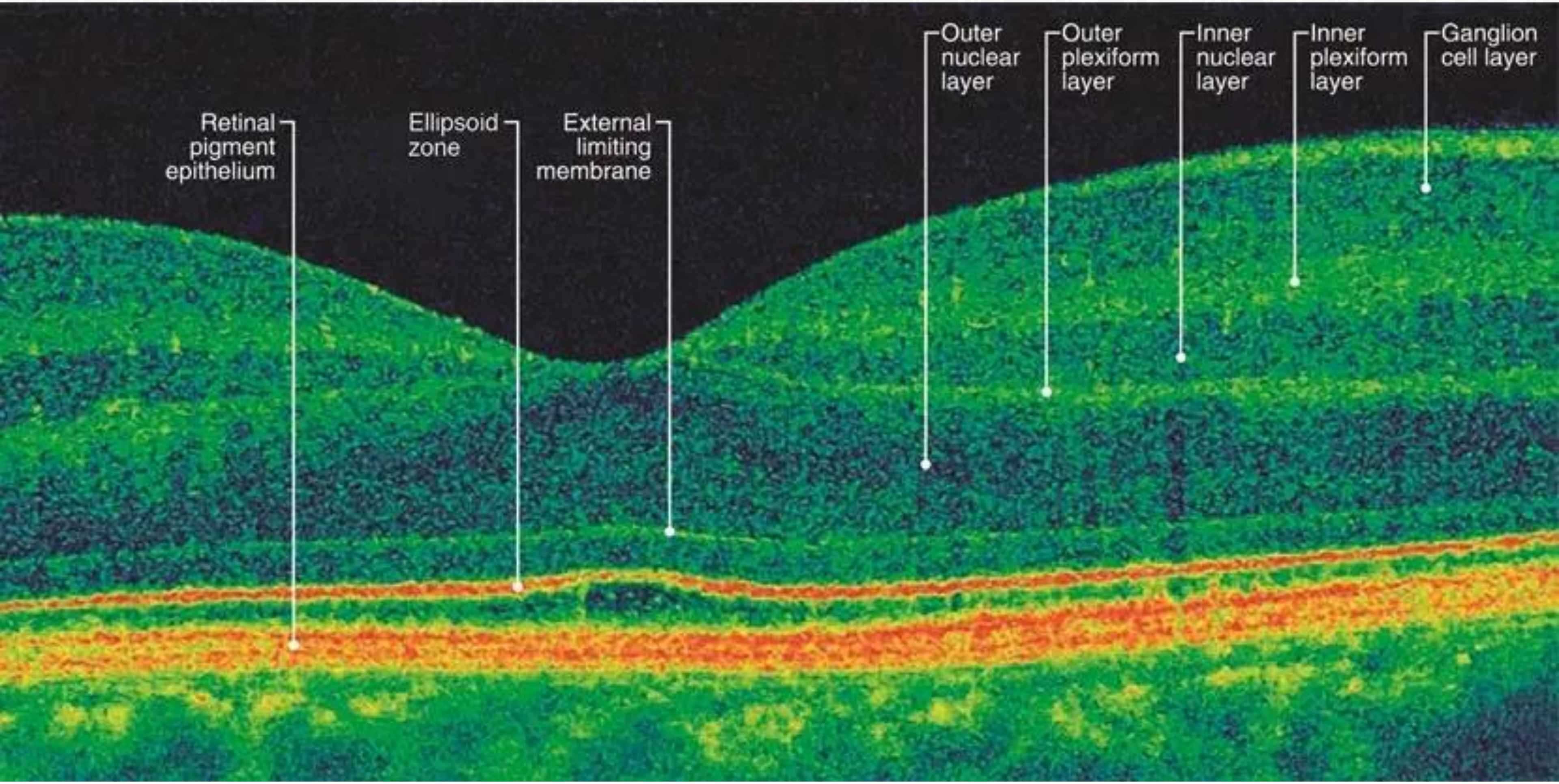
Principle:

- OCT uses low-coherence light (e.g. long-wavelength (near-infrared), broad-bandwidth light source) to measure the time delay and intensity of light reflected from retinal structures.

How It Works:

- 1. **Light Emission:** A light source emits a beam toward the retina.
- 2. **Reflection and Interference:** Light reflects back from different retinal layers.
- 3. **Detection:** An interferometer measures the light’s echo time and intensity.
- 4. **Image Construction:** Software reconstructs cross-sectional images.

Optical Coherence Tomography (OCT)



OCT Image of Retinal Layers

Image source: https://www.oscarwylee.com.au/glasses/eye/fluorescein-angiography

UCF Center for Research in Computer Vision

Optical Coherence Tomography (OCT)

Why It's Important:

- Detects microscopic changes in retinal layers early, often before symptoms appear.
- Essential for diagnosing, monitoring, and managing **retinal diseases** and other eye conditions.

Comparison with Other Modalities:

- **Compared to FA:** OCT is non-invasive and does not require dye injection.
- **Compared to Fundus Photography:** OCT provides **depth information** by visualizing individual retinal layers.

Applications of OCT

1. Diabetic Retinopathy (DR)

- **Use Case:**
  - Detects macular edema and thickened retinal layers.
- **AI Application:**
  - Automated segmentation of fluid pockets and retinal layers.

2. Age-Related Macular Degeneration (AMD)

- **Use Case:**
  - Identifies drusen, retinal atrophy, and subretinal fluid in wet AMD.
- **AI Application:**
  - Classifies AMD stages and predicts progression risk.

Applications of OCT

3. Glaucoma

- **Use Case:**
  - Measures the thickness of the retinal nerve fiber layer (RNFL) to detect optic nerve damage.
- **AI Application:**
  - Quantifies optic nerve damage and predicts disease progression.

Questions?

Surgical Visual Understanding

What is Surgical Visual Understanding?

Application of computer vision and machine learning techniques to analyze surgical videos.

- Key tasks:
  - Surgical procedure recognition
  - Instrument detection and tracking
  - Phase segmentation and workflow analysis
- Importance:
  - Enhancing surgical training, evaluation, and safety.
  - Providing real-time assistance to surgeons.

Video Modalities

1. Standard RGB Videos

- **Capture Method**: Captured using traditional cameras placed in the operating room or mounted on surgical lights.
- **Features**:
- Provides natural color information.
- Typically high-resolution (HD or 4K).
- Wide field of view, showing the entire operating room or procedure area.
- **Applications**:
- General surgical procedure recording for training, documentation, or analysis.
- Surgeon hand movement analysis.

Video Modalities

2. <u>Endoscopic Videos</u>

- **Capture Method**: Captured using an endoscope inserted into the patient's body during minimally invasive surgeries.
- **Features**:
- Close-up view of internal anatomy.
- High resolution and frame rate for real-time visualization.
- Limited field of view, focused on the surgical site.
- **Applications**:
- Instrument tracking and tissue analysis during minimally invasive procedures.
- Surgical phase recognition in procedures like laparoscopies.

Surgical Visual Understanding

3. Robotic Surgery Videos

- **Capture Method**: Captured using cameras integrated into robotic surgical systems (e.g., da Vinci Surgical System).
- **Features**:
- High-definition, stereoscopic views provide depth perception.
- Real-time synchronization with robotic arms.
- Controlled lighting and stable views of the surgical site.
- **Applications**:
- Instrument localization and tracking.
- Precise spatial reasoning and motion control for robotic-assisted surgeries.

Surgical Visual Understanding

4. <u>Depth Videos</u>

- **Capture Method**: Captured using stereo cameras or depth sensors.
- **Features**:
- Provides 3D structural information about the surgical environment.
- Often integrated with RGB data (RGB-D).
- Can detect instrument and tissue movements in 3D space.
- **Applications**:
- 3D modeling of the surgical field.
- Advanced spatial reasoning for instrument manipulation.

Surgical Visual Understanding - Applications

1. **Surgical Procedure Recognition**
- Automatic identification of surgical steps.
- Use in operating room monitoring and surgeon feedback.

2. **Instrument Detection and Tracking**
- Identifying and tracking surgical tools in real time.
- Avoiding errors and improving efficiency.

3. **Workflow Analysis**
- Segmentation of procedures into phases.
- Understanding surgical patterns for training.

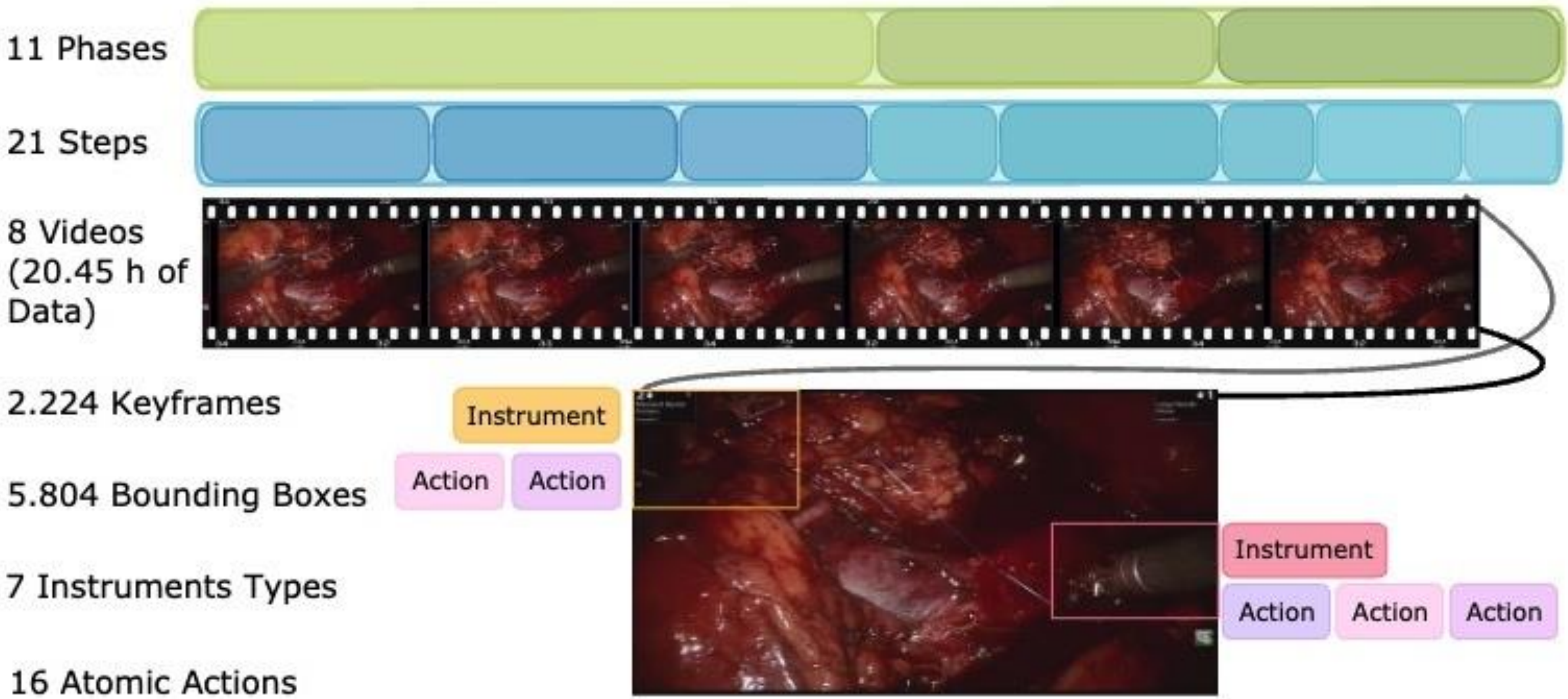
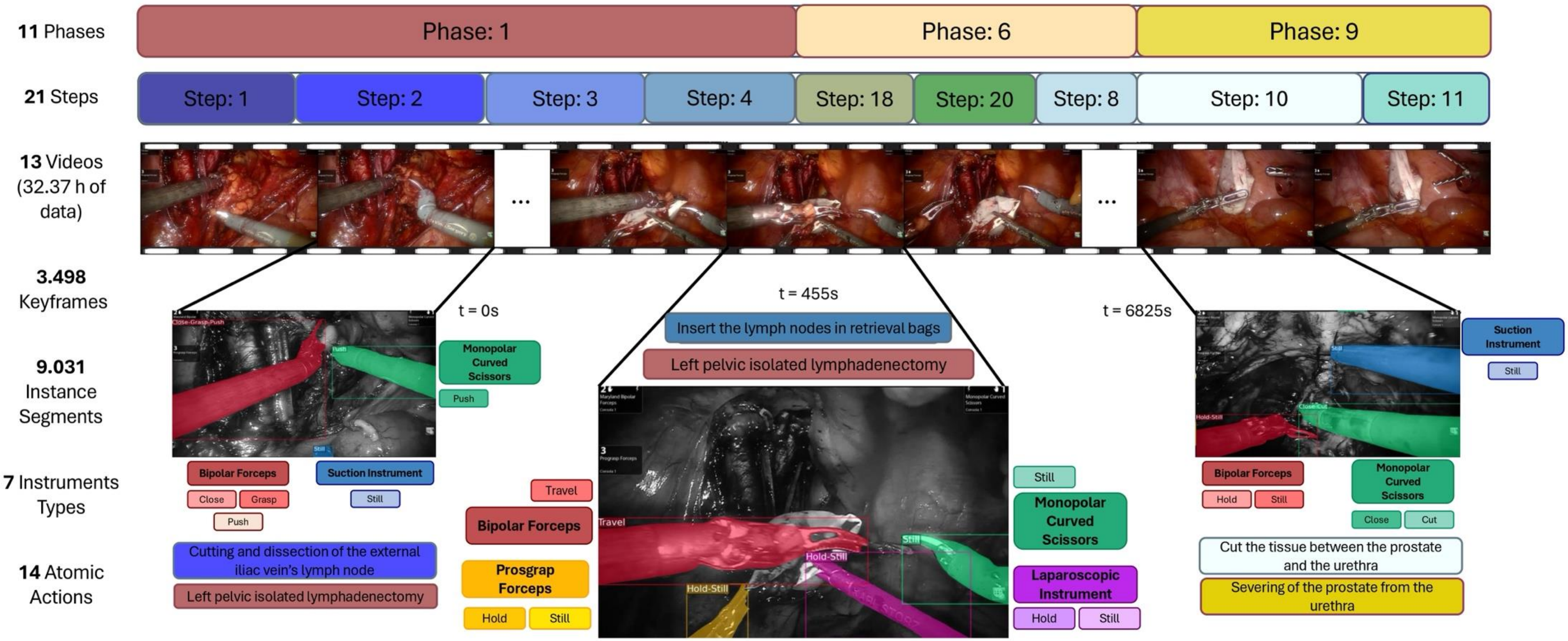


Fig. 1: **PSI-AVA Dataset** enables a holistic analysis of surgical videos through annotations for both long-term (Phase and Step Recognition) and short-term reasoning tasks (Instrument Detection and Atomic Action Recognition).

Valderrama, N., Ruiz Puentes, P., Hernández, I., Ayobi, N., Verlyck, M., Santander, J., ... & Arbeláez, P. (2022, September). Towards holistic surgical scene understanding. In International conference on medical image computing and computer-assisted intervention (pp. 442-452). Cham: Springer Nature Switzerland.



Ayobi, Nicolás, Santiago Rodríguez, Alejandra Pérez, Isabela Hernández, Nicolás Aparicio, Eugénie Dessevres, Sebastián Peña et al. "Pixel-Wise Recognition for Holistic Surgical Scene Understanding." arXiv preprint arXiv:2401.11174 (2024).

Other Signal Modalities in Medical/Healthcare

Beyond visual data, many medical/healthcare applications rely on signals like **EEG, ECG, EMG, PPG**, and others.

These signals are typically time-series data but can be transformed into image-like representations for certain analyses.

- **Key Modalities:**
- 1. **EEG** (Electroencephalogram) – Brain activity.
- 2. **ECG** (Electrocardiogram) – Heart activity.
- 3. **EMG** (Electromyogram) – Muscle activity.
- 4. **PPG** (Photoplethysmogram) – Blood volume changes.

Signal-to-Image Transformation

Why Transform Signals into Images?

- Enables the use of image-based machine learning techniques (e.g., CNNs, vision transformers).
- Facilitates better feature extraction, especially for frequency-based patterns.

Common Transformation Techniques:

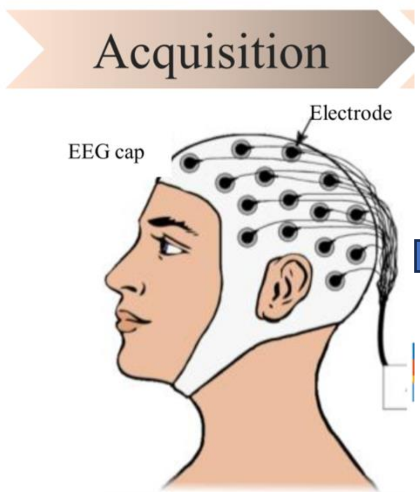
1. **Frequency Domain Analysis**:

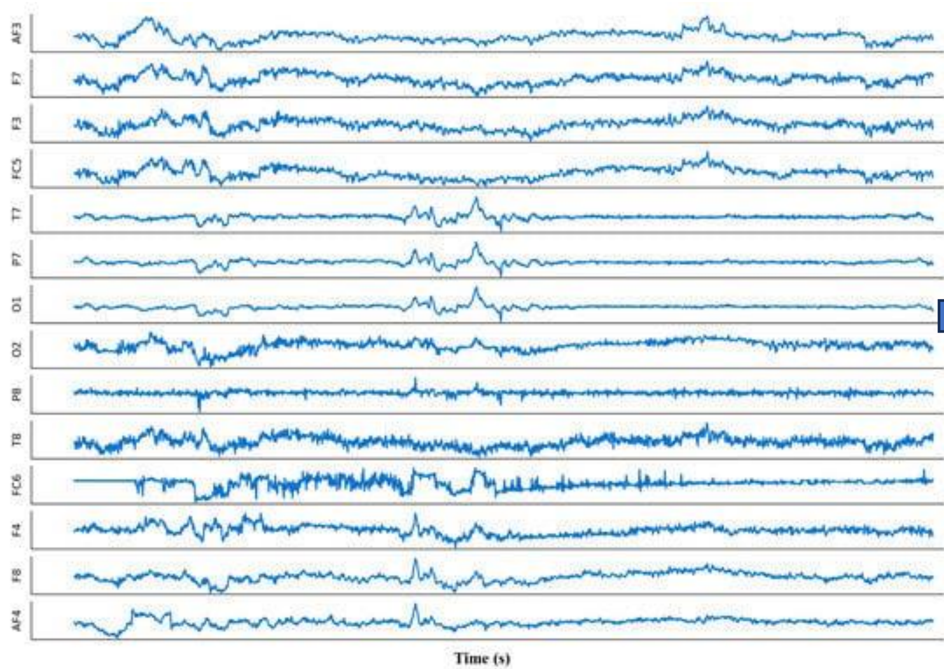
- Transform signals using techniques like Fourier Transform or Wavelet Transform.
- Represent signals as **spectrograms** or **scalograms**.

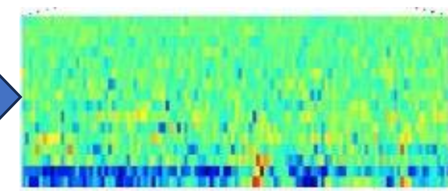
2. **Mel-Frequency Cepstral Coefficients (MFCCs):**

- Widely used in speech and sound processing.
- Converts time-domain signals into frequency-domain features.

Example: EEG







MFCC feature

Image source: Chaddad, Ahmad, Yihang Wu, Reem Kateb, and Ahmed Bouridane. "Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques." Sensors 23, no. 14 (2023): 6434.

Thank you!

Question?