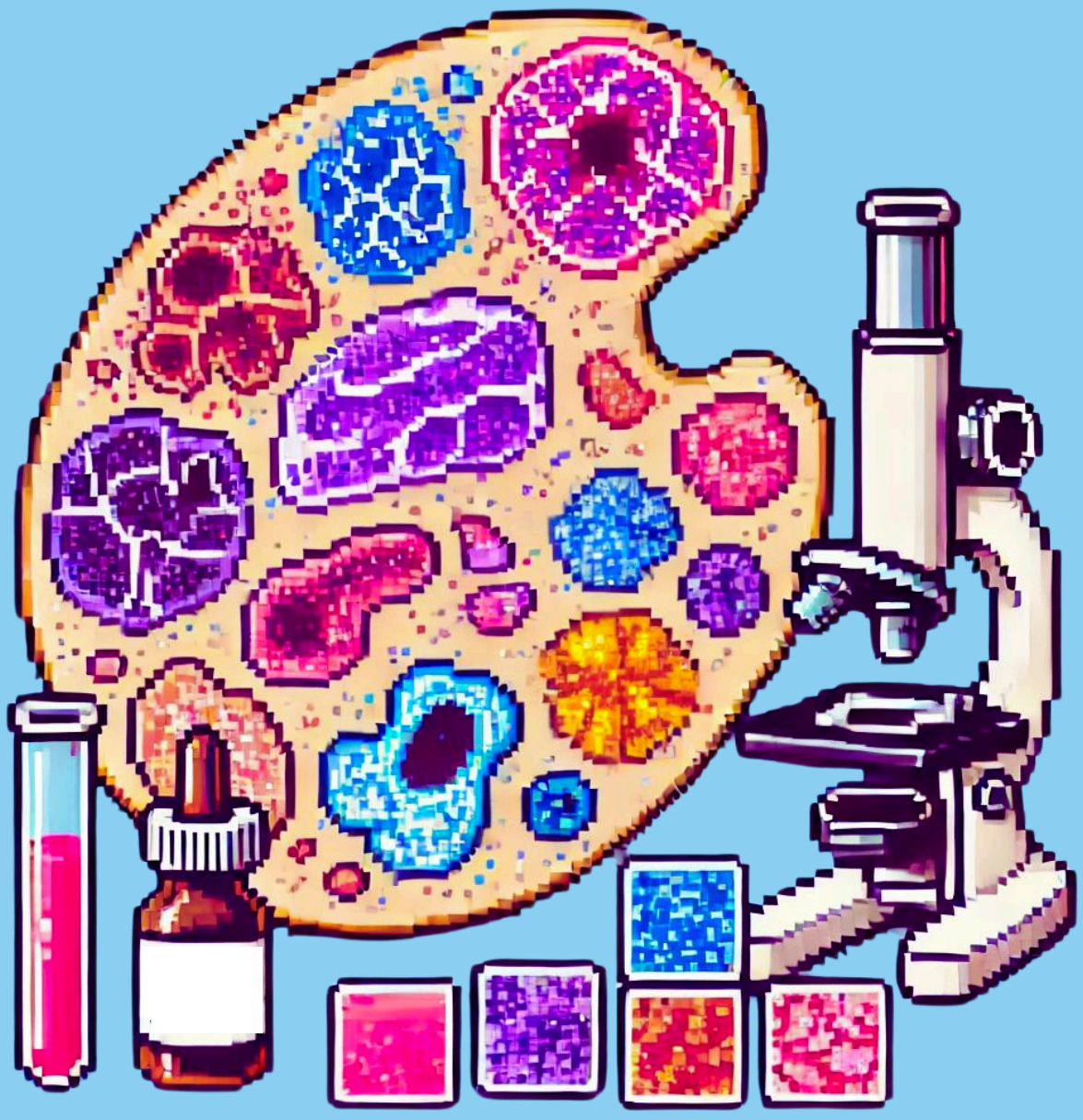
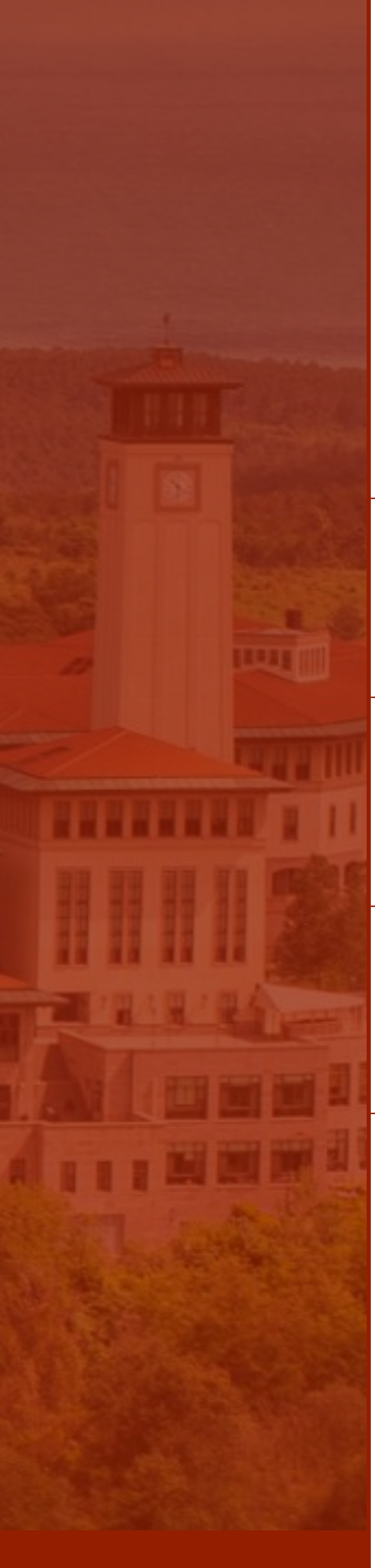
KUUJE

Koç University Undergraduate Journal of Engineering

Vol. 1, No. 2 | Summer 2024



Coloring the Canvas of Disease: The Art of Cellular Detail



Contents

Letter from the Editor-in-Chief

Dear Colleagues and Readers,

As we release the second issue of the Koc University Undergraduate Journal of Engineering, I want to reflect on the journey we've had in bringing this edition to fruition. Unlike our inaugural issue, where the excitement of creating something new fueled every step, this time has been about refining and sustaining our work. The transition from launching a new project to maintaining and improving it can be challenging and require difficult choices, and this issue was no exception.

The summer break posed its own set of difficulties. Coordinating with researchers, reaching out to reviewers, and managing our internal team dynamics during this period proved to be quite a task. However, I am pleased to share that we have successfully navigated these challenges and produced an issue that meets the exceptional quality we always strive for. This would not have been possible without our entire team's hard work and perseverance; for that, I am sincerely grateful. I would especially like to thank Ahmet Burak Kılıç for his exceptional organizational work and leadership in coordinating the works of our editorial board and all of its tasks.

This issue also marks a significant milestone for KUUJE as for the first time we had the pleasure of working with summer interns from TED Ankara College. These young minds that joined us in this academic endeavor, displayed exceptional work ethic and a genuine excitement to contribute. We hope that our collaboration has inspired them to embrace the world of science and consider becoming the next generation of engineers and scientists. For his contributions in helping us organize and establish the KUUJE Future Researchers Program, I'd like to sincerely thank Barış Şahin. With this collaboration, we hope to lay the groundwork for a future where we can contribute significantly to cultivating a culture of scientific passion from a young age and embedding this passion within the Turkish youth and academia.

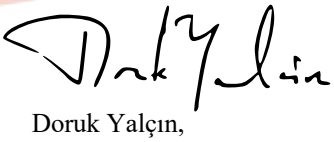
I am honored that this issue will be released on August 30th, a day of profound significance in Turkey. Victory Day marks a pivotal moment in our nation's history and symbolizes the triumph of independence and democracy and the beginning of modern Turkey. We owe a deep debt of gratitude to Ghazi Mustafa Kemal Atatürk and the countless individuals whose sacrifices and vision laid the foundation for the Republic. The reforms initiated by Atatürk transformed our nation not only politically but socially and created an environment where science and education could flourish. It is because of these enduring legacies that we are able to contribute to the global scientific community today.

On this day as we commemorate our history, let us also look forward with hope and determination, continuing to pursue knowledge and innovation with the same spirit of resilience and dedication.

I would like to thank all of the researchers who submitted their research to KUUJE, the reviewers who took the time to diligently go through all of the papers that we published, and my team members who worked day and night to make all of this a reality.

Thank you for your continued support. I look forward to your feedback and engagement with this latest issue and hope to talk to you all soon in our Fall issue.

Sincerely,



Doruk Yalçın,
Editor-in-Chief
Koc University Undergraduate Journal of Engineering

KUUJE

Koç University Undergraduate Journal of Engineering

MASTHEAD

Contact Information

Editor in Chief
editor_in_chief@kuuje.org
Editorial Board and Submissions
submit@kuuje.org
Sponsorships
sponsorships@kuuje.org
Corporate Relations
outreach@kuuje.org

Rumelifeneri Yolu 34450 Sarıyer,
İstanbul / Türkiye

EDITORIAL REVIEW

KUUJE
Koç University Undergraduate Journal of Engineering

A Complete Guide to Histopathological Staining and Imaging: Techniques and Applications

Doruk Yalçın[1]
Department of Computer Engineering
Faculty of Engineering
Koç University
Istanbul, Turkiye
dyalcin21@ku.edu.tr

Acar Karabıçak[2]
Department of Industrial Engineering
Faculty of Engineering
Bilkent University
Ankara, Turkiye
acar.karabicak@ug.bilkent.edu.tr

Osman Murat Doğu[3]
TED Ankara College
Ankara, Turkiye

***Abstract*—Histopathology relies heavily on accurate staining and imaging techniques to enable the detailed study of tissue diseases. This review comprehensively examines the dyes and staining methods that are integral parts of histopathological practices, tracing their historical development from traditional stains like Hematoxylin and Eosin to advanced fluorescent dyes, immunohistochemistry, and virtual staining. This review also looks at the chemical compositions, applications, and staining procedures of a wide range of dyes. Additionally, this paper also showcases the optical principles that govern the light-tissue interactions that are essential for optimizing various microscopy techniques. The paper further discusses the latest advances in microscopy, including brightfield, phase contrast, fluorescence, and super-resolution microscopy. This review aims to provide a valuable resource for researchers seeking to expand their understanding and application of histopathological techniques.**

Keywords—Histopathology, Tissue Staining, Microscopy Techniques, Hematoxylin and Eosin (H&E), Fluorescent dyes, Immunohistochemistry, Optical principles, Light-tissue interaction, Histological imaging, Super-resolution microscopy

I. Introduction

Histopathology, the microscopic study of tissue diseases, is a crucial part of modern medicine and allows for detailed insights into the structure and function of biological tissues. Accurate staining and visualization of tissue sections are essential for reliable diagnosis and research as they allow for the clear observation of cellular and tissue architecture. Over time, various staining techniques have been developed to enhance the contrast and specificity of histological observations, each designed to highlight particular components of cells and tissues.

This review provides a detailed examination of the dyes and staining methods that are fundamental to histopathological practices by first looking at the history and development of dyes, covering the progression from traditional stains like Hematoxylin and Eosin to more advanced methods involving fluorescent dyes and immunohistochemistry. This review also looks into the chemical compositions, applications, and staining procedures of a wide range of dyes and offers a comprehensive understanding of their roles in improving the visibility of cellular structures.

In addition, this paper addresses the optical principles that govern the interactions between light and biological tissues, which are essential for the optimizing of imaging techniques. The review covers various microscopy methods, including brightfield, phase contrast, fluorescence, and super-resolution microscopy, each with its own advantages and challenges in examining histological samples.

This review examines both traditional knowledge and recent advances in staining techniques and microscopy and highlights the importance of these tools in histopathology. This comprehensive examination aims to be a resource for both established and emerging researchers in the field and provide a solid foundation for further developments in histological staining and imaging.

II. History and Development of Dyes

A. Early Use of Natural Dyes

1) Ancient Mediterranean Civilizations: The use of natural dyes has deep historical roots, dating back to ancient civilizations in the Mediterranean. Among the most significant developments was the widespread use of Tyrian purple, a dye of great cultural and economic importance. Extracted from the mucus of Murex sea snails, Tyrian purple was first produced by the Phoenicians around the 15th century BCE. The complex and labor-intensive process required tens of thousands of snails to produce even a small amount of dye, making it extremely valuable. As a result, garments dyed with Tyrian purple became symbols of wealth, power, and prestige, often reserved for royalty and high-ranking officials.[1] In later centuries, the prestige of this dye was so profound that in the Byzantine Empire, the term 'born in the purple' (Latin: Porphyrogenitus) was used to describe imperial children born to reigning emperors specifically in the imperial palace's Tyrian purple dyed birthing room, signifying their inherent right to rule. [2]
The Phoenicians established trade networks throughout the Mediterranean, spreading the use of Tyrian purple across regions such as Egypt, Greece, and Rome. This dye was a luxury item but also a form of currency and a significant trade

Written: 16.08.2024 | Reviewed: 26.08.2024 | DOI: 10.5281/zenodo.13375858
Doruk Yalcin[1]: Undergraduate Student, College of Engineering, Koç University, Istanbul, Turkiye; Visiting Student Researcher at Ozcan Research Lab, University of California, Los Angeles, USA

commodity, contributing to the prosperity of Phoenician cities like Tyre and Sidon.

Image 1: Constantine VII Porphyrogenitus being baptized by Patriarch Nicholas Mystikos by Madrid Skylitzes

In addition to Tyrian purple, other natural dyes were also widely used in the Mediterranean region. The Greeks and Romans utilized dyes such as indigo, derived from the indigo plant, and madder, which produced a range of red hues. Saffron, extracted from the stigmas of the crocus flower, was another prized dye, known for its vibrant yellow color and used both in textiles and in religious rituals.

B. Discovery of Microscopic Life and Early Staining Techniques

1) The Discovery of Microscopic Organisms: The discovey of microscopic organisms started with Robert Hooke and Antoni van Leeuwenhoek, two Fellows of The Royal Society, who, between 1665 and 1683, discovered the first microscopic organisms. Hooke's "Micrographia" (1665) provided the first published depiction of a microorganism, the microfungus Mucor. Later, Leeuwenhoek observed and described protozoa and bacteria, a significant milestone in the study of biology.

2) Early Cell Discovery: In 1665, Robert Hooke further advanced the field by improving the design of the compound microscope, allowing him to observe cork cells. He coined the term "cells" due to their resemblance to monastic cells, marking the beginning of cellular biology.[3]

C. Development and Application of Early Stains

1) Carmin and Iodine Staining: Carmine, derived from cochineal insects, has been used historically to stain biological tissues, highlighting nuclei, chromosomes, or mucins depending on the formulation. However, challenges with dye quality were frequently noted. In contrast, iodine staining emerged as a key technique involving the complexing of iodine with crystalline compounds or polymers, leading to significant applications in detecting starch and distinguishing different tissue structures.[4][5]

2) Pioneering Histochemical Reactions: The early 19th century saw significant advancements in histochemical reactions. H.F. Link's discovery in 1807 of using iron sulfate to detect tannic acid in leaves, and M. Raspail's iodine-based microchemical detection of starch in 1825, resulted in more precise staining techniques. The Feulgen reaction, developed by R. Feulgen and H. Rossenbeck in 1924, became a cornerstone for DNA detection in histology.[6]

D. Introduction of Synthetic Dyes

The synthesis of the first artificial dye by William Henry Perkin in 1856 marked a transformative period in staining technology. Perkin's discovery of mauveine, derived from aniline, was initially a byproduct of his attempt to synthesize quinine. Recognizing the significance of his discovery, Perkin commercialized the dye, leading to the establishment of the synthetic dye industry, which quickly expanded to Germany and Switzerland. This innovation completely overhauled staining techniques in biology and medicine, providing a broader palette of colors.[7]

E. Fluorescent Dyes and the Evolution of Fluorescence Microscopy

1) The Observation of Fluorescence: Fluorescence, a phenomenon where certain substances emit light when exposed to specific wavelengths, was first observed by Sir Frederik William Herschel in 1845. His discovery that a quinine solution exhibited a blue color under sunlight led to further exploration by George Stokes, who described this effect in detail in 1852. The development of fluorescence microscopy, driven by the need for better differentiation in imaging, was significantly advanced by Helmholtz in 1874 and later expanded with the advent of fluorescent antibody labeling in the early 1940s.

2) Applications in Modern Microscopy: The use of fluorescent dyes has since become integral in labeling and detecting specific molecules within biological specimens. The cloning of Green Fluorescent Protein (GFP) in the early 1990s further enhanced the ability to visualize proteins within living cells, marking a significant evolution in microscopic techniques.

F. Immunochemistry and Autofluorescence

1) Immunohistochemistry: Immunohistochemistry (IHC) is as a technique for detecting specific antigens within tissue sections. Pioneered by Albert Hewett Coons and colleagues in 1941, the use of fluorescein isothiocyanate (FITC)-labeled antibodies allowed for the precise localization of antigens, revolutionizing diagnostic pathology and cancer research.[8]

2) Autofluorescence in Diagnostics: Autofluorescence, the intrinsic emission of light by certain cell and tissue components, has been recognized for nearly a century. Initially noted by Policard in 1924 for tumor diagnostics, the development of high-sensitivity imaging techniques in the late 20th century has renewed interest in autofluorescence for both research and clinical applications. This technology enables real-time, non-invasive examination of live cells and tissues, offering valuable insights into cell morphology and metabolism.

III. Traditional Histological Dyes

It is important to note that much of the content in Sections III. Traditional Histological Dyes and IV. Fluorescent Dyes is obtained from the *Handbook of Biological Dyes and Stains: Synthesis and Industrial Applications* by R.W. Sabnis. This

text provides a detailed exploration of the synthesis, properties, and applications of various dyes and stains used in biological research. For readers seeking more comprehensive information, this handbook serves as an excellent resource. [9]

A. Hematoxylin

1) Overview: Hematoxylin is one of the most fundamental stains in histopathological staining, particularly in nuclear staining. It is a natural dye derived from the logwood tree, *Haematoxylum campechianum*. Hematoxylin, when oxidized to hematein, forms a complex with metal ions such as aluminum, creating a dye that strongly binds to nucleic acids. This property makes it essential in the very widely used Hematoxylin and Eosin (H&E) stain.

2) Chemical Composition: Hematoxylin is a flavone compound with the molecular formula $C_{16}H_{14}O_6$ and a molecular weight of 302.28 g/mol. It appears as white to yellowish crystals that turn red upon exposure to light. The dye is soluble in water, ethanol, ethylene glycol, and methyl cellosolve. Hematoxylin itself is not a staining agent until it is oxidized to hematein, which then binds to metal salts like aluminum to create a strong nuclear stain. This complex imparts a blue to purple color to the nuclein in tissue sections, with a characteristic absorption maximum, λmax, at 292 nm.

3) Applications: Hematoxylin is primarily used in the staining of cell nuclei in various histopathological preparations. Its main application is in the H&E stain where it provides contrast to the eosin-stained cytoplasm and extracellular matrix. Hematoxylin is also used in the staining protocols of blood smears, collagen, neurons, and nucleic acids.

4) Staining Procedure: Hematoxylin staining procedure involves treating tissue sections with the dye after oxidation and mordanting. The stained sections are then differentiated to remove excess dye, followed by a counterstain such as eosin to highlight other tissue components.

B. Eosin Y

1) Overview: Eosin Y, also known as Eosin Yellowish, is an acidic dye that is primarily used as a counterstain in combination with Hematoxylin in the Hematoxylin and Eosin (H&E) stain which is the most commonly used staining technique in tissue diagnostics. Eosin Y selectively stains cytoplasmic components and extracellular matrix and provides a pink to red contrast against the blue or purple nuclei stained by Hematoxylin.

2) Chemical Composition: Eosin Y is a xanthene dye with the molecular formula $C_{20}H_6Br_4Na_2O_5$ and a molecular weight of 691.85 g/mol. It appears as red-brown crystals with non-fluorescent properties at lower pH and green fluorescence at pH 3.0. The dye binds to positively charged tissue components such as proteins in the cytoplasm by forming salt linkages with amino groups. It has a maximum absorption, λmax, at 517 nm and an emission peak at 538 nm though it is primarily used in non-fluorescent applications.

3) Applications: The primary application of Eosin Y is as a counterstain in the H&E stain where it highlights the cytoplasm, connective tissues, and extracellular matrix in shades of pink to red. This contrast allows for the clear differentiation of tissue structures, making it a fundamental part of histopathology. Eosin Y is also used to stain blood smears, bone marrow, and various cell components, including the cytoplasm and cell membranes. Additionally, it has applications in staining proteins and certain biological structures like keratin fibers.

4) Staining Procedure: Eosin Y is applied after the tissue sections have been stained with Hematoxylin. The sections are immersed in Eosin Y solution, followed by differentiation and dehydration steps, resulting in the characteristic pink to red colorization of the cytoplasm and extracellular matrix.

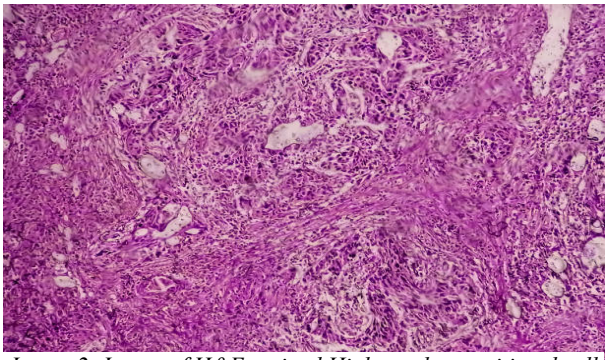


Image 2: Image of H&E-stained High-grade transitional cell carcinoma

C. Congo Red

1) Overview: Congo Red is an azo dye known for its application in detecting amyloid deposits in tissues. Discovered in the late 19th century, it was initially developed as a textile dye- where it was recently abandoned due to its carcinogenic properties- but quickly found its place in biological staining. Congo Red is unique in its exhibition of apple-green birefringence under polarized light when bound to amyloid fibrils.

2) Chemical Composition: Congo Red is an azo dye with the molecular formula $C_{32}H_{22}N_6Na_2O_6S_2$ and a molecular weight of 696.66 g/mol. It is a brownish-red powder soluble in water and ethanol. Congo Red's chemical structure consists of two azo groups (-N=N-) linking aromatic rings.

3) Applications: Congo Red is primarily used for the identification of amyloid deposits which are associated with diseases such as Alzheimer's and various systemic amyloidoses. The dye is applied to tissue sections and the presence of amyloid is confirmed by the characteristic apple-green birefringence under polarized light. Congo Red is also used for the staining of bacteria, collagen, fungi, and certain proteins involved in neurodegenerative diseases.

4) Staining Procedure: Tissue sections are stained with an aqueous solution of Congo Red, followed by differentiation in an alkaline alcohol solution. The stained sections are then examined under polarized light to detect amyloid deposits.

D. Alcian Blue

1) Overview: Alcian Blue 8GX is a cationic dye particularly valued for its ability to stain acidic polysaccharides such as glycosaminoglycans, mucosubstances, and mucins. It was developed in the mid-20th century and quickly became a standard tool for identifying and studying tissues with high acidic mucopolysaccharide content such as cartilage and mucus-secreting tissues.

2) Chemical Composition: Alcian Blue 8GX is a copper phthalocyanine dye with a molecular formula of $C_{56}H_{68}Cl_4CuN_{16}S_4$ and a molecular weight of 1298.93g/mol. Its chemical structure includes multiple cationic groups that interact with the acidic groups of polysaccharides and form stable complexes. The dye is soluble in water and ethanol and exhibits a maximum absorption, λmax, at 615nm- giving it a characteristic dark blue color.

3) Applications: The primary application of Alcian Blue 8GX in histopathology is in the detection of acidic mucosubstances and is particularly effective in staining glycosaminoglycans in connective tissues, mucus in goblet cells, and the cellular matrix. The dye is also used in hybridization assays, detection of hyaluronic acid, and other glycoprotein-rich tissues. Additionally, Alcian Blue has been applied in biological studies involving the locomotion of amoeba and quantitative determination of glycosaminoglycans.

4) Staining Procedure: Alcian Blue 8GX staining typically involves treating the tissue section with the dye under acidic conditions (pH 2.5 or lower). The acidic environment promotes the binding of the cationic dye to negatively charged acidic polysaccharides, resulting in a distinct blue coloration of target tissues.

E. Toluidine Blue O

1) Overview: Toluidine Blue O is a versatile basic dye belonging to the phenothiazine class, used for its metachromatic staining properties. First synthesized in the 19th century, Toluidine Blue is particularly useful in highlighting acidic tissue components such as glycosaminoglycans and nucleic acids. Its ability to stain different structures in distinct colors is noteworthy.

2) Chemical Composition: Toluidine Blue O has the molecular formula $C_{15}H_{16}N_3S$ and a molecular weight of 305.83 g/mol. The dye is typically available as a dark green powder, soluble in water and ethanol. It is characterized by two absorption peaks at 626 nm and 630 nm. The dye's monochromatic properties allow it to change color when binding to different tissue components: it stains nucleic acids and acidic tissues blue, and stains glycosaminoglycans in shades of purple or red. The dye's ability to interact with tissue components through electrostatic interactions and hydrogen bonding underlies its metachromatic behavior.

3) Applications: Toluidine Blue O is primarily used in the staining of acidic tissue components such as nucleic acids and glycosaminoglycans. It is especially effective in the highlighting of cartilage, mast cell granules, and mucins. The dye is used in various diagnostic applications such as Mohs' micrographic surgery for skin cancer where it helps delineate tumor margins. This stain can also be used in cytopathology for quick screening of cytological smears and in the detection of oral lesions.

4) Staining Procedure: Toluidine Blue O is applied to tissue sections in an aqueous solution. The dye selectively binds to acidic components, producing a color change that highlights the different structures within the tissue. The stained sections are then examined under a light microscope.

F. Crystal Violet

1) Overview: Crystal violet, also known as Gentian Violet, is a triphenylmethane dye, first synthesized in the late 19th century. It is one of the key components of the Gram stain which is a fundamental technique in bacterial classification.

2) Chemical Composition: Crystal violet has the molecular formula $C_{25}H_{30}ClN_3$ and a molecular weight of 407.98 g/mol. The dye typically appears as a dark green powder and is soluble in water, acetone, and chloroform; but, is insoluble in ether and xylene. It exhibits a pH-dependent color change, ranging from yellow at pH 0.0 to blue-violet at pH 2.0. The dye exhibits maximum absorption, λmax, at 590 nm. Crystal Violet binds strongly to cellular components through ionic and van der Waals interactions in the staining of bacteria, nucleic acids, and various other tissue structures.

3) Applications: Crystal violet is most commonly used in the Gram staining technique where it serves as the primary stain for differentiating Gram-positive from Gram-negative bacteria based on cell wall properties. In histopathology, Crystal Violet is used to stain cell nuclei, amyloid plaques, and certain proteins. It is also employed in staining skin, liposomes, and thrombocytes as well as vaginal smears for cytological analysis.

4) Staining Procedure: In the Gram stain procedure, Crystal Violet is applied to bacterial smears where it stains all cells initially. Following treatment with iodine, a mordant, the dye forms a complex that is retained in Gram-positive bacteria after decolorization, while Gram-negative bacteria lose the stain.

G. Basic Fuchsin

1) Overview: Basic Fuchsin is a triphenylmethane dye that is commonly used in the Ziehl-Neelsen stain for acid-fast bacteria, such as Mycobacterium tuberculosis. The dye also plays a key role in the Feulgen reaction which is used for detecting DNA. Its magenta color is popular for cases with need for intense contrast.

2) Chemical Composition: Basic Fuchsin has the molecular formula $C_{20}H_{20}ClN3$ and a molecular weight of 337.85 g/mol. It typically appears as dark green crystals that are soluble in water and ethanol but insoluble in ether. The dye's color is due to the presence of conjugated double bonds within its triphenylmethane structure and it has a maximum absorption, λmax, at 543 nm. Basic Fuchsin binds strongly to acidic tissue components and DNA.

3) Applications: Basic Fuchsin is primarily used in the staining of acid-fast bacteria through the Ziehl-Neelsen

method where it highlights mycobacterial cell walls in a bright magenta color. It is also used in the Feulgen reaction for DNA detection as well as in other histological stains where it can highlight nuclei, cytoplasm, collagen, and other tissue components. The dye is also used for staining fungi, bacteria, and certain proteins.

4) Staining Procedure: In the Ziehl-Neelsen staining technique, Basic Fuchsin is applied to heat-fixed bacterial smears, followed by decolorization and counterstaining, which results in the selective staining of acid-fast organisms. For the Feulgen reaction, Basic Fuchsin is used after hydrolysis of DNA, where it binds specifically to aldehyde groups formed in the DNA backbone.

H. Oil Red O

1) Overview: Oil Red O is an azo dye used for its ability to stain neutral lipids and fats in tissue sections. It has become a standard stain for identifying lipid accumulations, particularly in frozen tissue sections, ever since its introduction. The dye is extensively used in research relating to metabolic diseases such as atherosclerosis, obesity, and fatty liver disease.

2) Chemical Composition: Oil Red O has the molecular formula $C_{26}H_{24}N_4O$ and a molecular weight of 408.49 g/mol. It is a dark greenish-brown powder that is slightly soluble in water and ethanol but highly soluble in organic solvents such as acetone, benzene, and toluene. Oil Red O is a lipophilic dye, meaning that it has a high affinity for lipid-rich environments. It is an azo dye characterized by its chromophore structure which binds to neutral lipids and fats, producing a red stain. The dye has an absorption maxima, λmax, at 359 nm and 518 nm.

3) Applications: Oil Red O is primarily used for staining lipids in frozen tissue sections and is a cornerstone of studies relating to lipid metabolism and storage. It is effective in staining neutral lipids, lipid droplets, lipoproteins, and other fat-containing structures in tissues. The dye is frequently used in research involving aortic valves and arterial walls. It is also employed in studies of adipose tissue, myocellular lipid content, and metabolic disorders like fatty liver disease.

4) Staining Procedure: The staining procedure for Oil Red O involves preparing a working solution in an organic solvent which is then applied to frozen tissue sections. The sections are subsequently washed and counterstained with a water-soluble dye, allowing for the visualization of lipid-rich areas, which appear red under a light microscope.

I. Sudan Black B

1) Overview: Sudan Black B, also known as Solvent Black 3, is an azo dye used for the staining of lipids in biological tissue. Discovered in the 20th century, the dye has become commonplace in the identifying of fat-containing structures in cells and tissues. It is particularly useful in the diagnosis of hematological disorders such as acute myeloid leukemia, where it is used to detect lipids in leukocyte precursors.

2) Chemical Composition: Sudan Black B has the molecular formula $C_{29}H_{24}N_6$ and a molecular weight of 456.54 g/mol. The dye appears as a dark brown powder and is insoluble in water but soluble in organic solvents such as ethanol, acetone, benzene, toluene, and xylene. Sudan Black B is an azo dye characterized by its chromophore structure which allows it to strongly bind to lipids. The dye's absorption maxima, λmax, is at 415 nm and 598 nm- which corresponds to its dark coloration when bound to lipids.

3) Applications: Sudan Black B is widely used for staining lipid-rich tissues and cells. It is particularly effective for staining glycol methacrylate embedded tissue sections, blast cells, and lipids in various tissues. The dye is also used to stain lipoproteins, apolipoproteins, myelin, and neuron-specific nuclear protein NeuN. Its ability to stain lipids makes it a good tool for the diagnosis of lipid storage diseases and in the characterization of myeloid cells in acute myeloid leukemia.

4) Staining Procedure: The staining process with Sudan Black B involves applying the dye to tissue sections that have been prepared with organic solvent to retain lipids. The dye selectively binds to lipid droplets and membranes which results in the blue-black coloration. Stained sections are then examined under a light microscope to identify lipid-containing structures.

J. Aniline Blue

1) Overview: Aniline Blue, also known as Methyl Blue, is a triphenylmethane dye widely used commonly for the staining of connective tissue like collagen as well as cartilage and some types of microorganisms. Aniline Blue has applications in complex staining protocols such as Masson's trichrome stain which differentiates between muscle, collagen, and fibrin.

2) Chemical Composition: Aniline Blue is a triphenylmethane dye with a molecular formula of $C_{37}H_{27}N_3Na_2O_9S_3$ and a molecular weight of 799.80 g/mol. Its chemical structure includes sulfonated aromatic rings which allow it to bind strongly to the acidic components of the tissues, particularly collagen fibers. The dye is typically presented in dark blue or dark brown powder coloring and is soluble in water and ethanol but is insoluble in xylene. It absorbs light at a maximum wavelength, λmax, of 600nm- giving it a deep blue color in stained tissue.

3) Applications: Aniline Blue is primarily used for the staining of collagen fibers in connective tissues. It is commonly used in Masson's trichrome stain where it stains collagen blue, allowing for a sharp contrast against the red-stained muscle fibers. Additionally, Aniline Blue is used to stain cartilage, chitin, hyaline droplets, and certain fungal structures. It is utilized in biological assays for nucleic acids, protein determination, and the identification of Candida albicans.

4) Staining Procedure: Aniline Blue staining is achieved by applying the dye to the tissue sections after the initial staining in trichrome protocol.

IV. Fluorescent Dyes

A. DAPI

1) Overview: DAPI (4',6-diamidino-2-phenylindole) is a fluorescent dye used in the staining of DNA in cell nuclei. Originally developed for the detection of DNA in various biological samples, DAPI has since become a standard tool in histology for visualizing nuclear material. The dye binds preferentially to A-T rich regions in the minor groove of double-stranded DNA, producing a strong fluorescent signal that is easily detected under a fluorescent microscope. Due to its high specificity and bright blue fluorescence, DAPI is particularly useful in applications where accurate nuclear staining is required such as in the identification of apoptotic cells or in the evaluation of chromatin structure.

2) Chemical Composition: DAPI is an indole-based compound with the molecular formula $C_{16}H_{15}N_5$ and a molecular weight of 350.25 g/mol. It is a yellow powder in its physical form and is soluble in water and N,N-dimethylformamide, making it suitable for various protocols. The dye exhibits absorption maxima, λmax, at 358 nm and 342 nm, and when bound to DNA it emits fluorescence at 461 nm and 450 nm, which is the blue spectrum. The blue emission is a key feature that allows clear visualization of DNA in stained samples.

3) Applications: DAPI is used in both research and clinical settings to stain DNA in cells, bacteria, chromosomes, and microorganisms. Its applications include nuclear apoptosis assays, targeting drugs to cerebral neurons, and studying neurodegenerative diseases such as Alzheimer's disease and Down's syndrome. DAPI is also employed in monitoring bioaerosols and in the selective staining of nuclear material in various experimental models. Additionally, its ability to stain polynucleotides and proteins makes it a versatile tool.

4) Staining Procedure: The staining process with DAPI involves preparing the tissue or cell samples and applying the dye at a concentration optimized for the specific application. DAPI is usually added to cells fixed with paraformaldehyde or other fixatives, even though it can also penetrate live cells. After the application, the dye binds to the A-T rich regions of DNA and the excess dye is washed away. The stained samples are then examined under a fluorescence microscope where the bright blue fluorescence of the nuclei provides a clear contrast against the background.

B. Texas Red

1) Overview: Texas Red is a hydrophilic, red-emitting fluorophore derived from sulforhodamine 101 and is used in immunofluorescence and flow microfluorometry studies. Developed to complement fluorescein-based fluorophores, Texas Red emits a distinct red emission that does not overlap with the green emission of fluorescein, making it a good candidate for dual-label experiments. Texas Red conjugates maintain strong fluorescence and retain the biological activity of the proteins they label, which is crucial in the studying of cell surface markers, protein interactions, and cellular distribution.

2) Chemical Composition: Texas Red is a sulfonylchloride derivative of sulforhodamine 101, which is characterized by its hydrophilic nature due to the presence of sulfonic acid groups. It has the molecular formula $C_{31}H_{29}ClN_2O_6S_2$ and a molecular weight of 625.15 g/mol. The chemical structure of Texas Red facilitates its conjugation with amino groups on proteins and other biomolecules, leading to the formation of highly fluorescent

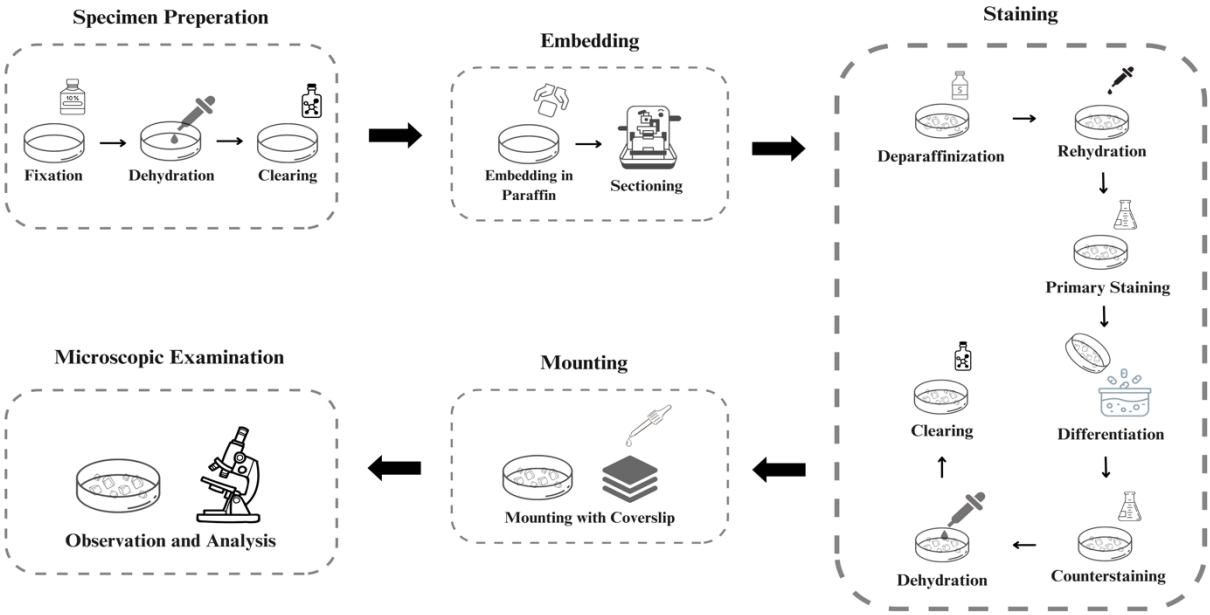


Figure 1: Steps for Biological Specimen Preparation and Examination

and water soluble conjugates. This dye has an excitation maximum of 596 nm and an emission maximum of 615 nm, making it compatible with fluorescence microscopy and flow cytometry systems equipped with the appropriate lasers. The hydrophilicity of Texas Red improves its solubility in aqueous solutions, reducing the likelihood of protein precipitation during labeling.[10]

3) Applications: Texas Red is commonly used in fluorescence microscopy and flow microfluorimetry studies, particularly in combination with fluorescein isothiocyanate (FITC). It has applications in the labeling of antibodies, lectins, and other proteins. For example, Texas Red conjugates have been employed to study Fc receptor distribution on various cell populations. Texas Red's resistance to photobleaching compared to other fluorophores like fluorescein and tetramethylrhodamine makes it suitable for long-term imaging studies.

4) Staining Procedure: The staining of Texas Red involves covalent attachment of the dye to proteins through its sulfonylchloride group which reacts with free amino groups on the target proteins. This conjugation is typically performed at a pH of around 9.0, where the reaction is highest. The process is carried out at low temperatures (0°C) to minimize hydrolysis of the dye and maximize protein labeling. After conjugation, unbound Texas Red is removed through techniques such as gel filtration or dialysis. The resulting Texas Red-labeled proteins can then be used in various fluorescence-based assays, where their bright red fluorescence can be detected independently of other labels, such as FITC, providing clear and distinct signals for multiple targets within a single experiment.

C. FITC

1) Overview: Fluorescein Isothiocyanate (FITC) was introduced in the mid-20th century, and has been widely adopted for labeling proteins, antibodies, and other biomolecules due to its bright green fluorescence and ease of conjugation. FITC is favored for its water solubility, high quantum yield, and minimal nonspecific binding to biological tissues. These properties make it an ideal dye for applications such as immunofluorescence, flow cytometry, and confocal microscopy, where precise detection of specific biomolecules is required.

2) Chemical Composition: FITC is a xanthene dye with the chemical formula $C_{21}H_{11}NO_5S$ and a molar mass of 389.38 g/mol. It is composed of a fluorescein core structure modified by an isothiocyanate group, which enables its covalent attachment to proteins and other molecules through thiourea linkages. This chemical structure allows FITC to produce a strong fluorescence signal when excited by blue light, emitting primarily in the green to yellow spectrum. FITC's excitation maximum is at approximately 488 nm, and its emission maximum is around 517 nm, making it compatible with standard fluorescence microscopy and flow cytometry systems.

3) Applications: FITC is used in a variety of biological assays such as immunofluorescence microscopy, where it is conjugated to antibodies to visualize the distribution of antigens within cells and tissues. It is also used in flow cytometry for the analysis of cell populations based on the expression of surface markers. Additionally, FITC is used in labeling various proteins, nucleotides, and other biomolecules for tracking and quantification in different experimental setups. However, FITC is sensitive to pH changes, photobleaching, and overlap with cellular autofluorescence; these problems can be mitigated by using antifade reagents and optimizing experimental conditions.[11]

4) Staining Procedure: The staining process with FITC involves conjugating the dye to the target protein or antibody through its isothiocyanate group. This reaction is carried out in a slightly alkaline buffer, usually around pH 9.5, where the isothiocyanate group reacts with the amine groups of the protein, forming a stable thiourea bond. Typically, the reaction is performed at 4°C in the dark to prevent photobleaching. After the conjugation, unbound FITC is removed through gel filtration or other purification techniques to ensure that only the labeled proteins remain. The resulting FITC-labeled proteins can then be used in a variety of fluorescence-based assays, providing a bright and reliable signal for detecting specific biomolecules in biological samples.

D. Cy3 and Cy5

1) Overview: Cy3 and Cy5 are two closely related cyanine dyes used for fluorescent labeling and are designed for conjugation to biomolecules like antibodies, DNA, and proteins, providing observable and non-overlapping emission spectra.

2) Chemical Composition: Cy3 is a sulfoindocyanine dye with a polymethine chain of three conjugated methine carbons, while Cy5 has a longer polymethine chain with five conjugated methine carbons. The chemical structure of Cy3 and Cy5 includes sulfonate groups that increase their water solubility, and succinimidyl ester groups that allow their conjugation to primary amines. Cy3 has an absorption maximum, λmax, at 550 nm and an emission maximum at 565 nm, while Cy5 absorbs maximally at 650 nm and emits at 667 nm. [12]

3) Applications: Cy3 and Cy5 are used in flow cytometry, fluorescence microscopy, and fluorescence in situ hybridization (FISH). Cy3 is typically used with excitation sources in the green spectrum, while Cy5 is used with red and near-infrared excitation sources. Their distinct spectral properties make them good choices for dual-label experiments, allowing simultaneous detection of multiple targets.

4) Staining Procedure: The staining process with Cy3 and Cy5 involves conjugation to target biomolecules using their succinimidyl ester groups, which react with primary amines on proteins or other targets. The reaction typically occurs in a slightly basic buffer, around pH 8-9, to facilitate efficient conjugation. After the conjugation, unreacted dye is removed through purification methods like gel filtration.

E. Rhodamine 6G

1) Overview: Rhodamine 6G is a xanthene dye known for its strong fluorescent properties, primarily emitting in the yellow-orange spectrum. It is commonly used for staining mitochondria, as well as various biological membranes and proteins.

2) Chemical Composition: Rhodamine 6G has the molecular formula $C_{28}H_{31}N_2O_3Cl$ and a molar mass of 479.02 g/mol. It appears as a red-brown or maroon powder or crystals and is highly soluble in water, ethanol, and methanol. Rhodamine 6G has an absorption maximum, λmax, at 528 nm and an emission maximum at 551 nm. The chemical structure includes an ethyl ester group allowing for its solubility and reactivity with various biological substrates.

3) Applications: Rhodamine 6G is used to stain mitochondria in live cells and fixed tissue samples due to its ability to selectively bind to these organelles. It is also used in measuring membrane potential, detecting nucleic acids, and in apoptosis and protease activity detection.

4) Staining Procedure: The staining process with Rhodamine 6G involves preparing the dye in a such as water or ethanol, and then applying it to the biological sample. The dye readily permeates cells and selectively stains mitochondria and other membranes. After staining, excess dye is washed away, and the sample is examined under a fluorescence microscope using the appropriate filters to detect the characteristic yellow-orange fluorescence of Rhodamine 6G.

V. Immunohistochemistry in Histopathology

A. Principles in Immunohistochemistry

Immunohistochemistry (IHC) is a technique used in histopathology used to visualize the presence and localization of specific antigens within tissue sections. Utilizing the specificity of antigen-antibody interactions using IHC, enables the detailed mapping of protein expressions, giving insight into cellular and molecular processes. [13]

The core principle of IHC revolves around the binding of antibodies to their corresponding antigens, which are typically proteins, within a tissue sample. This binding is then visualized using various detection methods that amplify the signal, allowing pathologists and researchers to observe where and how much of the antigen is present. [14]

1) Antigen-Antibody Binding: Antibodies, which are Y-shaped proteins, are designed to bind to specific epitopes- unique molecular signatures- on antigens. These antibodies can be either monoclonal, derived from a single cell clone and highly specific to one epitope, or polyclonal, which are derived from multiple cell lines and can bind to multiple epitopes on the same antigen, offering greater sensitivity but less specificity.

2) Sample Preparation and Fixation: For IHC to be effective, the tissue samples must be properly prepared. This process involves fixation, which preserves tissue architecture and antigenicity by cross-linking proteins. The most common fixative used is 10% neutral buffered formalin (NBF). However, the fixation process must be carefully controlled; overfixation can mask antigenic sites, while underfixation can lead to tissue degradation and loss of antigen.

3) Antigen Retrieval: Fixation can also create cross-links between proteins that mask antigenic sites. To unmask these sites and allow antibodies to access their target epitopes, antigen retrieval techniques are employed. The most widely used method, Heat-Induced Epitope Retrieval (HIER), involves heating the tissue sections in a buffer solution to break the cross-links. Alternatively, enzymatic retrieval methods can be used, though these are typically reserved for specific antigens that are sensitive to heat.

4) Blocking Non-Specific Binding: Before applying the primary antibody, it is crucial to block any non-specific binding sites within the tissue to reduce background staining. This is typically achieved using normal serum or other protein-based blocking solutions that prevent the antibodies from binding to non-target sites. Additional blocking strategies may include the use of synthetic peptides or protein mixtures that are designed to mask non-specific sites. This step ensures that the staining observed is specific to the antigen of interest.

5) Detection Systems: Once the primary antibody has been bound to its antigen, the next step is to visualize this interaction; this step is conducted with detection systems. The most common method is the use of a secondary antibody that binds to the primary antibody. This secondary antibody is usually conjugated to an enzyme, such as horseradish peroxidase (HRP) or alkaline phosphatase (AP), which catalyzes a color-producing reaction when exposed to the appropriate substrate. In some cases, fluorescence-based detection systems are used where the secondary antibody is conjugated to a fluorophore, allowing visualization under a fluorescence microscope.

6) Visualization and Interpretation: The final step of IHC is the visualization of the stained tissue. The reaction between the enzyme and its substrate produces a color change at the site of antigen-antibody binding, which can be observed under a light microscope. The intensity and pattern of staining provides valuable information about the distribution and quantity of the antigen within the tissue. This process allows pathologists to diagnose diseases, identify tumor types, and assess the expression of biomarkers.

B. Chemical Components of Immunohistochemistry

1) Fixatives: Fixatives are vital in IHC for preserving the structural integrity of tissue samples and maintaining antigenicity. The choice of fixative can significantly influence the outcome of an IHC experiment. Formalin-based fixatives, such as 10% neutral buffered formalin, are commonly used due to their ability to create stable cross-links between proteins. These cross-links effectively lock tissue structures in place, preventing autolysis and preserving antigens for subsequent detection. Alternatives to formalin such as alcohol-based fixatives or acetone are occasionally used particularly in situations where rapid fixation is needed, such as in frozen sections.[14]

2) Antibodies: The efficacy of an IHC assay is largely dependent on the selection and validation of antibodies.

Monoclonal antibodies are ideal for applications requiring consistent and reproducible results as they offer high specificity by recognizing a single epitope on the antigen. In contrast, polyclonal antibodies provide greater sensitivity but with a potential trade-off in specificity due to cross-reactivity. Proper validation of antibodies, including titration and testing against known controls is essential to ensure that they produce clear and specific staining with minimal background.[15]

3) Chromogens: Chromogens are chemical compounds that produce a color precipitate when reacted with an enzyme in the detection system. The selection of chromogen is crucial as it determines the visual contract and clarity of the IHC results. Diaminobenzidine is a commonly used chromogen that produces a stable brown precipitate which is easily visible against a hematoxylin counterstain. Other chromogens such as 3-amino-9-ethylcarbazole (AEC) produce different colors which allow for flexibility in multiplex staining protocols where multiple antigens need to be visualized simultaneously. The stability, intensity, and compatibility of the chosen chromogen with the detection system and tissue type are important factors which must be considered to obtain optimal staining results.

C. *Applications in Diagnostic Pathology*

Immunohistochemistry is a widely utilized technique across various fields of medicine and a crucial tool for the identification and localization of specific cellular antigens in tissue sections. While its applications extend beyond oncology to areas such as infectious diseases, neurodegenerative disorders, and autoimmune conditions, in this section, we'll focus on its application in cancer to provide a complete picture of its impact and relevance in a single branch of medicine.

In oncological pathology, the diagnosis of invasive ductal carcinoma of the breast relies not only on the identification of characteristic morphological features but also on the expression of key biomarkers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor as detected by IHC. These biomarkers are crucial for determining the appropriate therapeutic regimen, including the use of hormone therapy or HER2-targeted therapies.[16]

Histopathology combined with IHC is very effective in the classification of lymphomas. The differentiation between Hodgkin and non-Hodgkin lymphoma as well as the subclassification of non-Hodgkin lymphomas into various types such as diffuse large B-cell lymphoma or follicular lymphoma, relies heavily on the pattern of antigen expression. Markers such as CD20, CD3, CD30, and PAX5 are routinely assessed using IHC to classify these lymphomas accurately, which directly influences the treatment plan including chemotherapy regimens and targeted therapies such as rituximab.

Techniques such as fluorescence in situ hybridization (FISH), polymerase chain reaction (PCR), and next generation sequencing (NGS) are increasingly being integrated with histopathological analysis to identify genetic mutations, gene fusions, and other molecular alterations that have diagnostic, prognostic, and therapeutic implications. For example, in non-small cell lung cancer (NSCLC), histopathological analysis combined with molecular testing for mutations in epidermal growth factor receptor, anaplastic lymphoma kinase, c-ros oncogene 1, and v-raf murine sarcoma viral oncogene homolog B1 has become routine practice. The detection of these mutations not only confirms the diagnosis but also identifies patients who are likely to benefit from targeted therapies such as tyrosine kinase inhibitors (TKIs). Similarly, in colorectal cancer, the detection of KRAS, NRAS, and BRAF mutations informs the use of epidermal growth factor receptor inhibitors.

Histopathology is also used in the identification of prognostic and predictive biomarkers which are essential for patient stratification and treatment selection. The expression of certain proteins, as determined by IHC, can provide information on the likely course of the disease and the patient's response to specific therapies.

In breast cancer, the assessment of Ki-67, a marker or proliferation, provides prognostic information that can influence decisions regarding the intensity of chemotherapy. Similarly, the detection of mismatch repair (MMR) proteins in colorectal cancer, which can be lost due to mutations in the MMR genes, not only has prognostic implications but also predicts the response to immune checkpoint inhibitors, a class of immunotherapy.

In addition, the role of histopathology in identifying molecular alterations that drive tumorigenesis is expanding with the development of new biomarkers. The use of IHC to detect PD-L1 expression in tumors is now standard practice in selecting patients for immunotherapy with checkpoint inhibitors across various cancer types, including melanoma, lung cancer, and urothelial carcinoma.

VI. Optical Principles and Interactions

The interaction of light and biological matter and tissue follows the same fundamentals of optics. The interactions of absorption, transmission, reflection, refraction, and scattering are used for the optimizing of imaging techniques. The usage of these interactions follows a series of understandings and assumptions that make the evaluation of biological tissue possible in the process of histological staining and imaging. [17]

A. *Interaction of Light with Biological Matter*

1) Absorption: Photon absorption and electron excitation occurs at different rates in different molecules of the tissue. When photons are absorbed within the tissue and electron transitions and excitation occurs, the different absorption spectra of molecules causes the exhibition of distinguishable characteristic properties. This property can be exploited in histological staining with the usage of dyes. For example Hematoxylin, which binds to the cell nucleus and nucleic acids, emits a blue-purple coloration that highlights the nuclei within the cytoplasmic background. [18], [19]

2) Transmission: Transmission is the passage of light through a given medium. Biological tissue is a translucent structure and thus can't transmit light sufficiently. The insufficiency results in the necessity of the usage of stains to enhance visualization. The degree at which light is transmitted through the tissue can indicate the density and composition of the tissue; the transmission of light in stained tissues can be used in a manner that highlights the

morphological details of cellular structures. For example, it is possible through H&E staining to clearly differentiate dense nuclear material from the less-dense cytoplasmic regions. [20]

3) Reflection: Reflection is the process in which light is deflected from the tissue image surface. The reflection of light in samples is defining as different biological tissues have varying reflective properties depending on their surface structures and refractive indices. Compared to other optical properties, reflection is less prominent in the evaluation of traditional histological staining but is crucial in examining surface topologies using techniques such as scanning electron microscopy (SEM).

4) Refraction: Refraction is the bending of light through its transition between two mediums that have different refractive indices. As biological tissue is a heterogeneous structure with diverse composition and hydration levels, it contains a range of refractive indices. Through Abbe's equation, aberrations in tissue refraction can greatly reduce image quality and resolution. Immersion oils that have refractive indices similar to those of glass are commonly used to mitigate the aforementioned problem, which enhance resolution and image clarity at high magnification rates.

5) Scattering: Scattering occurs when light is redirected at several angles due to the heterogeneity within the tissue. Biological tissue is made up of a connection of arbitrarily shaped structures with different lengths, densities, sizes, and compositions. The interconnected nature of the cell makes it so that the structure can be evaluated on a continuously changing function. Phase contrast microscopy makes use of the scattering to visualize unstained cells in detail, converting phase differences into recognizable intensity variations. [20], [21]

6) Biological Tissue Considerations: Examining biological tissue requires numerous considerations as tissue is made up of various cell types, extracellular matrix components, and differing contents of water; these variables all influence the interaction with light. Components such as collagen, which can be found in the extracellular matrix, exhibit birefringence that can be visualized and observed with polarized light microscopy. This insight provides additional structural information. The water content in tissues has effects on the tissue's refractive index. Proper preparations of samples- employing fixation, dehydration, and mounting, ensures uniform light-tissue interaction and minimizes unwanted artifacts.

7) Magnification, Resolution, and Contrast:

a) Magnification :Magnification is the visual enlargement of an object. Magnification is achieved through the usage of lenses, including objective and ocular, in the optical microscope. Objective lenses are responsible for the initial magnification of the subject and are placed closest to the subject. Objective lenses are labeled by their magnification rate (e.g., 10x, 20x, 40x) and the numerical aperture (NA) which refers to the range of angles at which the lens can emit light. NA can be determined by the equation:

$$NA = nsin\theta,$$

where n is the medium's refractive index and is the half angle of the maximum cone of light that enters the lens. Ocular lenses are secondary magnification lenses that further magnify the visualization provided by the objective lens. Standard ocular lenses usually magnify the image by an additional rate of x10. While magnification is crucial for better observation of the specimen, it is also paramount to not lose image clarity.

b) Resolution: Resolution is the ability to distinguish two closely positioned points in the same image entity, the minimum resolvable distance. The resolution of the image depends on the wavelength of the imaging radiation and the numerical aperture. This is described in Abbe's equation:

$$d = \frac{0.612\lambda}{NA},$$

where d is resolution, λ is the wavelength of the imaging radiation, and NA is the numerical aperture

c) Contrast: Contrast is the observable difference in intensity between the subject object and the background in which it is captured. In histology, proficient observation of contrast is essential for the visualization of different cellular and tissue components. Certain staining techniques provide ease in the observation of different components by way of increased contrast; common dyes such as Hematoxylin and Eosin (H&E) provide contrast between nuclei and the cytoplasmic elements. Phase Contrast microscopy converts the phase shifts present in the light passing through the specimens into variations in intensity which provides the necessary contrast in unstrained samples. Differential Interference Contrast employs the usage of polarized light, which are made up of waves that vibrate in a single plane, to produce high contrast images of transparent specimens, highlighting the infinitesimal differences in the existing refractive indices. Dark-field microscopy uses oblique illumination, causing only scattered light to be collected which establishes contrast by making the specimen appear brighter in a dark background.

VII. Types of Microscopy

Microscopy is a fundamental part of histology and enables our examination of tissue, providing a tool for diagnosis and later treatment. As a long standing and ever developing area of study, various microscopy techniques have been developed and employed for numerous situations and needs. While some microscopy techniques have proved to be foundational and lasting, new microscopy techniques have been growing in popularity. As discussed in the previous section, the usage of essential optical interactions such as absorption, transmission, reflection, refraction, and scattering are pivotal properties used in the optimal elucidation of morphological and structural features.

A. Brightfield Microscopy

Brightfield Microscopy is often considered the most fundamental optical microscopy technique. The usage of Bright-field microscopy can be described as the most straight-forward method for optical illumination and morphological analysis of the characteristics of cells and tissue.

Bright-field microscopy operates by simply transmitting light through the specimen where through absorption of light, the specimen appears visually darker compared to the background.

1) Principles of Brightfield Microscopy:

a) Illumination: The specimen examined is illuminated from below by a halogen or LED lamp. The light is then focused on the specimen by using a condenser lens.

b) Image Formation: When the light is focused on the specimen, it is differentially absorbed by the components that make up the specimen; the differing absorption rates creates contrast which allows for the visualization of structures within the specimen. The objective lens, tasked with collecting the transmitting light, magnifies the image which is then further enlarged by the ocular lens.

c) Staining: As mentioned before, cell tissue and biological samples are often inherently transparent; this property results in a lack of sufficient contrast in the imaging to allow for clear observation. Specimens are stained with chemical dyes that bind to the cellular structure to enhance contrast. The most popular of the dyes used is Hematoxylin and Eosin (H&E) but simple stains such as Methylthioninium chloride and hexamethyl pararosaniline chloride can also be used when examining certain biological specimens and cell tissues.

2) Color Deconvolution and Error Estimation: Color deconvolution is a necessary step when evaluating stained biological samples. As dyes bind to the different components of the cell tissue, it produces a varying range of colors and contrast; color deconvolution allows us to properly separate and quantify the visual output of the stained specimen.[22]

a) Challenges in Color Deconvolution: The necessity for color deconvolution stems from the usage of multiple stains with overlapping spectral absorption properties. A composition of different colors that represent differing and specific biological structures are revealed when tissue samples are stained; accurately separating these colors is essential for the correct evaluation of the specimen examined. The assumption that the resulting stain concentration is linearly dependent on the absorption rate only holds under monochromatic light conditions and in practice, with the usage of polychromatic light, imaging often occurs with non-linear absorbance characteristics and significant deconvolution errors.

b) Advances in Deconvolution Techniques: Haub and Meckel address the complications of color deconvolution by using a mathematical model that simulates non-monochromatic absorbance values that provide a more accurate representation of typical imaging conditions. Their findings show that traditional linear deconvolution methods under polychromatic illumination can result in errors up to ten times higher than those using sequential monochromatic LED illumination. [22]

3) Usage of Brightfield Microscopy: Brightfield microscopy is best utilized for the relatively simple morphological analysis of stained tissue. It performs best when used for cases of routine histological examinations in the analysis of fixed and stained specimens. Brightfield microscopy is highly compatible with a variety of traditional staining techniques such as Hematoxylin and Eosin (H&E) for general tissue architecture observation, Gram staining for differentiating bacterial species, Periodic Acid-Schiff (PAS) staining for detecting polysaccharides such as glycogen, and Giemsa staining for blood smears and bone marrow samples. The general compatibility of Brightfield microscopy with several types of stains establishes the technique as a fundamental part of microscopy in general. [23]

B. Phase Contrast Microscopy

Phase contrast microscopy was introduced by Frits Zernike in 1934. The technique transforms phase shifts in light passing through transparent specimens into variations in light intensity, making otherwise invisible structures visible. This technique uses the wave nature of light and the differences in the created optical path length as light passes through specimens with varying thickness and refractive indices.[24]

1) Principles of Phase Contrast Microscopy:

a) Annular Illumination: Phase contrast microscopy uses the annular diaphragm in the condenser which creates a hollow cone of light that illuminates the specimen. As light passes through the different parts of the specimen, it undergoes phase shifts that depend on the specimen's optical properties.

b) Phase Plate: The, now phase shifted, light passes through a phase plate located in the back focal plane of the objective lens. This plate alters the phase of the unscattered light by a quarter wavelength, creating a constructive or destructive interference with the scattered light which has already been phase shifted by the specimen.

c) Interference and Contrast: Varying and visible shades of light and dark appear in the final image because of the light intensity caused by the interference between the scattered and unscattered light. This converts phase differences into amplitude differences, enhancing contrast and making structures visible without the need of staining.

2) High Resolution Techniques and Machine Learning Integration:

a) High-Resolution Phase Contrast Microscopy: Apodized phase contrast (APC) is a method that modifies the phase plate to reduce halo artifacts that are bright or dark rings that obscure details in traditional phase contrast images. Shimasaki et al. describe how APC techniques can enhance the visualization of subcellular structures to provide clearer and more accurate images with less distortion. [25]

b) Machine Learning Integration: In cases of low contrast in phase contrast imaging, the challenge of manually analyzing the morphological structures of cells can be effectively addressed using machine learning models. Using techniques such as deep convolutional neural networks, Yakovlev et al. were able to train a model using labeled datasets that was able to automatically recognize and segment cellular structures such as astrocytes.[22], [26]

3) Usage of Phase Contrast Microscopy: Phase contrast microscopy is mainly utilized in cases where it would be preferable to not stain the specimen to avoid damages or distortions to the cellular structures. Phase contrast microscopy allows for the cell's morphological structure to be observed clearly without using dyes or stains and is ideal for the observation of processes such as mitosis, motility, and any intracellular dynamics and interactions in real-time. This

technique is used extensively in the studying of bacteria, protozoa, and other microorganisms in their natural, unstained state. It is also employed in histological analysis where thin tissue samples with low natural contrast may not be suitable for observation with techniques such as brightfield microscopy.[25]

C. Fluorescence Microscopy

1) Principles of Fluorescence Microscopy: Fluorescence microscopy utilizes the principle of fluorescence where specific molecules known as fluorophores absorb light at a certain wavelength and emit it at a longer wavelength. This phenomenon is characterized by the Stokes shift where the emitted light has less energy than the absorbed light due to the energy loss during the transition. Common fluorophores include Fluorescein Isothiocyanate (FITC) which is excited by blue light and emits green fluorescence and DAPI (4',6-diamidino-2-phenylindole) which binds strongly to DNA and emits blue fluorescence when excited by ultraviolet light. [27]

2) Microscope Design and Light Path: In fluorescence microscopy, a laser or high-intensity lamp provides the excitation light; this light is filtered through an excitation filter that selects the specific wavelength required to excite the chosen fluorophore. The usage of different light wavelengths corresponding to specific dyes and their usages are: [28][29]

- FITC is excited by blue light and emits green fluorescence; FITC is used to label proteins and antibodies.
- Texas Red is excited by green light and emits red fluorescence, Texas Red is used to label secondary antibodies and helps visualize specific proteins and antigens in tissue samples.
- Cy3 is excited by green light and emits red fluorescence while Cy5 is excited by red light, emitting far-red fluorescence. Both Cy3 and Cy5 are used for the labeling of nucleic acids.

3) Usage Resolution and Two Photon Excitation: To achieve better imaging and higher resolutions in tissues, fluorescence microscopy often uses two-photon excitation. Two photon excitation is the process in which dyes such as Cy5 are excited by two photons of lower energy that results in simultaneous absorption of the photons and an emission of higher energy. This technique leads to deeper tissue penetration with reduced photodamage which is especially beneficial in the imaging of living cells and will be discussed in detail in the Multiphoton Microscopy section.

4) Usage of Fluorescence Microscopy: Fluorescence microscopy is beneficial in cases where high specificity and sensitivity are required to effectively visualize the morphological structures of cells and subcellular structures. This technique allows for the simple targeting of specific proteins, nucleic acids, and other cellular components using fluorophores which can easily be adjusted for the labeling target.[30]

D. Confocal Microscopy

1) Principles of Confocal Microscopy: Confocal microscopy is a non-invasive imaging technique that outputs high-resolution, three dimensional images of unprocessed, non-fixed tissue. Confocal microscopy employs point illumination and a spatial pinhole to eliminate out-of-focus light which improves image resolution and contrast. This technique effectively slices through the specimen tissue optically and creates a cross-sectional image by scanning the specimen with a laser and capturing the emitted light through a pinhole. Confocal microscopy, similar to other microscopy techniques, employs the usage of stains and can also be integrated with the usage of fluorophores.[27], [31], [32]

2) Fluorescence Confocal Microscopy: Confocal microscopy can employ the usage of fluorochromes such as Acridine Orange to increase the contrast between the different tissue components inside the specimen and highlight the morphological structure of the cell. A laser emitting at 488 nm (blue-fluorescence) excites the fluorophores inside the tissue; the emitted light is captured after passing through the spatial pinhole where the resulting image with improved resolution and contrast is obtained. The use of fluorophores outputs a high-quality image with a resolution comparable to traditional H&E stains.[33]

3) Usage of Confocal Microscopy: Confocal microscopy is best utilized in cases where swift, on-site diagnosis is needed. Unlike frozen and processed sections, which can introduce artifacts that compromise the integrity of the samples, confocal microscopy outputs high resolution imaging without freezing or cutting. This technique is employed extensively in cases where the original specimen may be needed for further analysis, reducing the need for additional biopsies and surgical interventions for additional surgical interventions.[34]

E. Multiphoton Microscopy

1) Principles of Multiphoton Microscopy: Multiphoton microscopy (MPM) is a technique in which two or more photons of lower energy are simultaneously absorbed by the fluorophore, resulting in imaging with relatively less photodamage. MPM uses longer wavelengths such as infrared to achieve deeper tissue penetration; this method deviates from the traditional single-photon confocal microscopy which uses a single, high-energy photon that may in some cases cause significant phototoxicity and has limited penetration depth compared to MPM.[35] In the attempt of achieving deep tissue imaging, light scattering becomes a significant issue that reduces image clarity; to prevent the loss of image clarity, optical clearing techniques such as benzyl alcohol/ benzyl benzoate (BABB) is used to match the refractive index of the tissue with the imaging medium which significantly reduces scattering. When combined with MPM and clearing, this allows for imaging depths of up to 1mm or more with a resolution rivaling traditional histology methods. This combination has been particularly effective in the generating of high quality, three dimensional images of tissues such as kidney biopsies and allowed for the visualization of glomerular structures and collagen matrix

organization which is not easily labeled in standard 2D histology [36].

2) Data Processing Challenges: As tissue depth increases the resolution of the images tend to decrease due to scattering. This results in a necessary trade off relationship between tissue depth and image resolution. The high resolution three dimensional images generated by MPM also have the added drawback of producing large amounts of data that requires computational tools for processing and interpretation that is not only large in volume but also complex in nature. Machine learning models have increasingly been integrated into the process of data interpretation to ease the processing of these large datasets.

3) Usage of Multiphoton Microscopy: MPM allows for the imaging of intact, unprocessed tissue samples in three dimensions which allows for a more complete visualization of the sample architecture; this three dimensional viewing reduces the likelihood of missing critical lesions which might be overlooked in traditional two dimensional visualization. In many clinical studies MPM is explored for its potential to replace or augment traditional histology; especially in renal pathology and gastroenterology the potential for multiphoton microscopy techniques can allow for significantly better imaging of biopsy samples with MPM being used to perform optical biopsies during endoscopy.[36], [37]

F. Electron Microscopy

1) Principles of Electron Microscopy: Electron Microscopy (EM) is a technique in which electrons are used for imaging instead of light, aiming to create a high resolution visualization due to the short wavelength of electrons. The primary types of electron microscopy are Transmission Electron Microscopy (TEM), Scanning Electron Microscopy (SEM), and variations such as Scanning Transmission Electron Microscopy (STEM) and Cryo-Electron Microscopy (Cryo-EM). [38]

a) Transmission Electron Microscopy: TEM has the electrons pass through an ultra-thin (<100 nm) section of the specimen. The interaction between the electrons and the sample provides information about the internal structure of the specimen down to an atomic level.

b) Scanning Electron Microscopy: SEM involves scanning the surface of a specimen with a focused electron beam. The electrons interact with the atoms on the surface, producing signals that provide information about the sample's topography and composition.

c) Energy-Dispersive X-ray Spectrometry (EDX) in SEM: EDX is often combined with SEM to provide elemental analysis of the sample by detecting X-ray emitted from the sample during electron beam exposure.

d) Field Emission Scanning Electron Microscopy (FE-SEM): FE-SEM is a variation of SEM that uses a field emission gun to produce high-resolution images with minimal sample damage.

2) Challenges and Innovations: Traditional electron microscopy techniques require extensive sample preparation and processing such as dehydration, embedding in resin, and the thin sectioning of samples; such procedures have the risk of damaging the sample and introducing artifacts into the imaging. Kawasaki et al. have proposed the NanoSuit method which allows for the examination of paraffin-embedded sections in a nondestructive manner using correlative light and electron microscopy (CLEM). The NanoSuit method encases the sample in a thin, dry, and vacuum-proof membrane that prevents dehydration and structural damage during SEM observation. This method preserves the integrity of the sample in imaging and is particularly beneficial in histological studies where the preservation of fine structures is critical.[39]

3) Usage of Electron Microscopy: EM is best utilized in histological cases in which specimens are required to be imaged in an ultra-high resolution with detailed visualization of cellular structures at the molecular level. Techniques such as transmission electron microscopy are especially useful when observing structures such as organelles or the arrangement of proteins. SEM is favored when the observation of detailed surface morphology is required and tissue topography is being studied.

G. Super Resolution Microscopy

1) Principles of Super Resolution Microscopy: Super resolution microscopy, which was recognized with the Nobel Prize in Chemistry in 2014, refers to a set of imaging techniques that break the limit of traditional light microscopy, allowing for imaging at a nanometer-scale resolution.[40]
The resolution achieved in super resolution microscopy is bound to the Point Spread Function (PSF), which describes how a point source of light is imaged by the optical system. The PSF can be mathematically represented as:

$$\mathrm{PSF}(f,z) = I_r(0,z,f)\exp\left[-z\alpha(f) - \frac{2\rho^2}{0.36\frac{cka}{\mathrm{NA}f}\sqrt{1+\left(\frac{2\ln 2}{c\pi}\left(\frac{\mathrm{NA}}{0.56k}\right)^2 fz\right)^2}}\right]$$

,where a is an adjustment factor, c is the speed of light, k-factor depends on the truncation ratio and level of irradiance, f is the photon frequency of the imaging beam, I_r is the intensity of the reference beam, NA is the numerical aperture, and ρ is the radial position from the center of the beam on the corresponding z-plane.

a) Photoactivated Localization Microscopy (PALM) and Stochastic Optical Reconstruction Microscopy (STORM): PALM and STORM achieve super-resolution by sequentially activating and then precisely localizing individual fluorescent molecules by fitting the point spread function (PSF) of each molecule to determine its position with high accuracy- typically in a range of 2 to 25nm. Aggregating these localized positions allows for observation of a super-resolution image with molecular-level detail.

b) Structured Illumination Microscopy (SIM): SIM uses patterned illumination to encode high resolution information into images. The sample is illuminated with a grid pattern where the resulting images are computationally reconstructed to achieve resolution beyond the diffraction

limit. SIM is particularly useful for imaging live cells with minimal photodamage [41].

c) Stimulated Emission Depletion (STED) Microscopy: STED employs dual lasers to deplete fluorescence in a controlled manner, shrinking the PSF and allowing for resolution improvements in the scales of tens of nanometers.

2) Challenges and Solutions:

a) Phototoxicity and Photobleaching: The high intensity illumination required for super-resolution imaging can lead to phototoxicity and photobleaching, especially in living-cell samples. Recent studies have proposed more photostable fluorophores and optimized imaging protocols to minimize the aforementioned damage. Techniques such as STORM and PALM use low-intensity light to sequentially activate fluorophores and reduce overall light exposure. [40]

b) Speed and Temporal Resolution: Capturing dynamic processes in live cells requires the use of high-speed imaging which is difficult due to the sequential nature of PALM and STORM. The usage of techniques such as Reflected Light Sheet Microscopy (RLSM) allow for selective plane illumination, reducing out-of-focus fluorescence and improving the signal-to-background ratio (SBR), enabling fast and accurate imaging of single molecules in living cells. [42]

c) Localization Accuracy and Image Reconstruction: Achieving high localization accuracy is necessary for super-resolution microscopy. The precision of molecule localization depends on factors such as the number of detected photons and the quality of the PSF fitting. Improving the computational algorithms and the development of better localization methods have significantly enhanced the resolution and accuracy of super-resolution images[43].

3) Usage of Super-Resolution Microscopy: Super-resolution microscopy allows for the studying of cellular structures in a way that is not resolvable with conventional light microscopy. PALM and STORM are ideal in the studying of the distribution and organization of proteins within cells at the molecular level, allowing for the visualization of the arrangement of proteins within cellular membranes or the organization of cytoskeletal filaments. SIM is well suited for live-cell imaging where minimal photodamage is required such as vesicle trafficking in real-time enhanced resolution [41]. STED microscopy is needed to study fine structures such as synaptic connections, nuclear pores, or protein complexes; the technique's ability to achieve resolution down to tens of nanometers makes it invaluable to the structural analysis of fine samples[40]. RLSM is crucial in situations where reducing out-of-focus light and improving SBR is essential. This technique is used commonly in the observation of transcription factors and other proteins within the nucleus. [42]

VIII. Image Analysis and Quantification

Image analysis for histopathology, in recent years, have evolved from a set of techniques handcrafted and tailored for feature extraction in certain datasets to a more sophisticated, automated approach that integrates deep learning into the analysis of images. This section explores the methodologies for multi-scale feature extraction- both traditional and novel- in obtained images, the integration of advanced imaging techniques such as cone-beam computed tomography and multimodal imaging, and the limitations and future directions of the field of image analysis in histopathology.

A. Traditional Methods of Feature Extraction and Their Limitations

Image analysis in histopathology has long been conducted with the usage of manual or semi-automated methods of feature extraction to identify specific features within histopathological images such as texture, shape, color, and structural characteristics. These features are then used to analyze and classify tissue samples; however, despite the historical significance and continued use of the traditional methods for image analysis, they are not without their limitations.

1) Handcrafted Features: Handcrafted features in traditional image analysis are manually designed based on domain knowledge and are tailor to capture specific aspects of tissue morphology. This technique can include texture analysis through the quantification of pixel intensities within a region, thereby allowing for the characterization of tissue patterns such as glandular structures or stromal architecture.

The use of Gabor filters allows for the analysis of the frequency and orientation of textures within an image. These filters have been effective in identifying repetitive structures such as the regular patterns found in certain types of epithelial tissues. Similarly, wavelet transforms have been used to decompose images into different frequency components and allow for the multi-scale analysis of tissue textures. These techniques provide information regarding the structural composition of tissue and aid in tasks such as tumor detection and grading. [44]

Handcrafted features, however, require significant domain expertise as features require tailoring to the specific characteristics of the tissue being analyzed. This can lead to a lack of generalizability as features that perform well on a given tissue may not perform well on another.

2) Energy-Based Methods and Active Contours: Energy based methods, including active contours or “snakes”, include the optimization of an energy function that guides the segmentation of objects within an image. For example active contour models are used to delineate the boundaries of structures such as cell nuclei by evolving a contour to minimize the energy associated with deviations from the object boundary. While these tools are powerful tools for segmentation they are limitingly sensitive to initialization and parameter settings. If the initial contour is not close enough to the object boundary or if the parameters are not optimized the segmentation may fail, leading to inaccuracies or incomplete delineations. In addition, these methods often struggle with images that contain significant noise or irregular structures, which are commonplace in histopathological samples. [45]

3) Feature Selection and Dimensionality Reduction: Feature selection aims to identify the most informative

features of a potentially large set, reducing the computational burden and interpretability of the model. Models such as Sequential Forward Selection (SFS) and Sequential Backwards Selections (SBS) have been used to iteratively add or remove features based on their contributions to classification performance.

Dimensionality reduction techniques such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are commonly used. PCA, for example, transforms the original features into a set of uncorrelated components that capture the most variance in data, reducing the dimensionality while retaining as much information as possible. LDA on the other hand, seeks to maximize the separation between classes by projecting data onto a lower-dimensional space where class distinctions are most apparent.[46]

4) Limitations and Challenges: Traditional methods of feature extraction are being increasingly overshadowed by their limitations which are inherent to their design. These methods that rely heavily on manual intervention and domain-specific knowledge are being surpassed by the advent of deep learning and more sophisticated computational models that offer automated, accurate, and generalizable approaches to histopathological images.

B. Multi-scale Feature Extraction

The multi-scale approach to feature extraction mimics the way pathologists analyze tissue samples at various magnifications, starting from a low resolution to capture broad structural patterns and then moving to higher resolutions to examine details. Multi-scale feature extraction is used to observe the complex and heterogeneous nature of histopathological images and allows for the visualization of structures with varying sizes and shapes.

1) Importance of Multi-scale Analysis: Histopathological images are inherently multi-scale and they contain information that varies in different levels of magnification. As magnification increases detailed features such as the arrangement of cells, nuclei, and subcellular structures become visible, which necessitates an analysis approach that can meaningfully extract features at multiple scales. One approach to this is to decompose images into multi-resolution representations using the Gaussian pyramid method which works by iteratively down-sampling an image and smoothing it with a Gaussian filter, creating a sequence of images with progressively lower resolutions. This method is useful for capturing broad structural patterns at lower resolutions and refining details as the analysis moves to higher resolution. This method allows for the system to focus on regions that require more detailed information for accurate classification. This approach has proven to be effective in the classification of whole-slide images where the system assigns labels to image tiles based on confidence levels achieved at each resolution. [44]

2) Techniques for Multi-scale Feature Extraction: Wavelet transforms is a common approach to multi-scale feature extraction; this method decomposes an image into components that capture information at different scales and orientations. Wavelet-based methods are effectively used in the analysis of textural patterns as they can capture both the coarse and fine details of tissue structures. [47]

Another technique is the use of multi-resolution pyramids such as the Laplacian pyramid. This technique progressively reduces the resolution of an image while preserving the essential features of the image, enabling the system to capture both the global structure and local details of the tissue. In the context of prostate cancer detection, Sridhar et al. demonstrated the effectiveness of a hierarchical classifier that operates at multiple scales. The classifier begins by analyzing the image at a lower magnification to discard non-tumor regions, thereby reducing the computational load at higher magnifications. As the analysis proceeds to higher scales, the classifier can focus on the remaining pixels which are more likely to be tumors. [48]

3) Applications of Multi-scale Feature Extraction: Multi-scale feature extraction has been significant for the improvement of diagnostic accuracy, particularly in the detection and classification of tumors. In prostate cancer, multi-scale feature extraction approaches have been shown to effectively differentiate between tumor and non-tumor regions by analyzing the textural and architectural patterns at different scales; this process is made less computationally burdensome by narrowing down the regions of interest before conducting detailed analysis.

In the classification of colorectal cancer and adenocarcinoma, this technique has proven to be effective in improving the performance of machine learning models. Multi-feature feature extraction has also been used to predict gene mutations and immune-related gene expression, demonstrating a high level of versatility in its applications.

C. Deep Learning for Histopathological Imaging

1) Convolutional Neural Networks (CNNs): Convolutional neural networks are an integral part of the recent developments in histopathological image analysis. CNNs are designed to automatically and adaptively learn spatial hierarchies of features from images through backpropagation, which is an algorithm that updates the model's weights by calculating the gradient of the loss function with respect to each weight and iteratively adjusting them to minimize the error. The ability to learn from raw data without the need for handcrafted features has made CNNs particularly well-suited for histopathology where the complexity and variability of tissue images can range significantly. Li et al. have used pre-trained CNNs on large datasets such as ImageNet and fine-tuned them on specific histopathological tasks such as tissue classification, gene mutation prediction, and other critical tasks. The hierarchical structure of CNNs allow for them to capture both low-level features such as edges and textures, and high-level features such as specific tissue types. [49][50][51]

2) Transfer Learning and Fine Tuning: A key innovation in applying deep learning to histopathology is the use of transfer learning, where a model pre-trained on a large dataset is adapted to a specific task with a smaller, domain-specific dataset. This is particularly beneficial in cases where obtaining large, annotated datasets can be challenging. Li et al. demonstrate this effectiveness by fine-tuning FTX-2048 an ImageNet-pretrained CNN; the fine-tuned model significantly outperformed the original pre-trained model in tasks such as tissue classification and gene expression prediction. The study found that fine-tuning allowed the model to adapt to the specific characteristics of the samples and led to improved feature extraction and classification accuracy. [49]

3) Advanced Architectures: Recurrent Neural Networks and Generative Models: CNNs have acted as the primary architecture used in histopathological image analysis but other advanced deep learning architectures such as recurrent neural networks (RNNs) and generative adversarial networks (GANs) are also gaining traction. RNNs, particularly Long Short-Term Memory (LSTM) networks, are well-suited for tasks that involve sequential data or where spatial dependencies are critical. RNNs have been explored for tasks such as mitosis detection and the analysis of tissue architecture. [52][53]

GANs have also introduced new possibilities for data augmentation and synthetic image generation. GANs consist of two neural networks, the generator and the discriminator, that are trained simultaneously in a competitive manner. The generator creates synthetic images while the discriminator evaluates those images' authenticity against real images. This process allows for the GANs to generate highly realistic histopathological images which can augment training datasets.

Boktor et al. utilized GANs for virtual histological staining and showcased the potential of GANs to replace traditional staining methods with computational techniques. Their multi-channel GAN approach allowed for the generation of high-quality stained images from unstained photon absorption remote sensing (PARS) images, reducing the need for chemical staining in histopathological analysis. Pillar et al. have also worked on the development of deep learning based virtual staining of nonfixed tissue histology. Their approach focuses on rapidly generating H&E-like images from nonfixed tissue which could significantly shorten the diagnostic turnaround time in surgical pathology by offering a faster alternative to conventional staining with high diagnostic accuracy.[54][55]

D. Integration of Advanced Imaging Techniques in Histopathology

The integration of advanced imaging techniques into histopathology has increased diagnostic accuracy by providing detailed, three dimensional, and multimodal insights into tissue structures. Techniques such as cone-beam computed tomography (CBCT) offer enhanced visualization and analysis capabilities that complement traditional histological methods. The combination of these imaging modalities with histopathological evaluation allows clinicians for more precise and comprehensive understanding of tissue structures.

1) Cone-Beam Computed Tomography (CBCT): CBCT provides three dimensional imaging capabilities that allows for a more accurate and comprehensive assessment of anatomical structures compared to traditional two dimensional imaging methods such as conventional radiography. In the study by Morelos et al. , CBCT was used to evaluate internal dental resorption- a condition characterized by progressive destruction of intraradicular dentin. The authors demonstrated that CBCT significantly improved the detection and localization of resorptive lesions within teeth, providing a more precise assessment of the extent of damage. This three dimensional visualization allowed for the identification of resorption lacunae and other structural abnormalities that weren't easily detected with traditional imaging methods.[56]

2) Multimodal Imaging and Digital Pathology: Multimodal imaging involves combining different imaging modalities, integrating the structural, functional, and molecular information displayed in different types of imaging methods. For instance, the combination of CBCT with magnetic resonance imaging (MRI) or positron emission tomography (PET) can provide both anatomical and functional information. This integration is particularly valuable in oncology where it is crucial to understand both the structural aspects of the tumor and its metabolic activity. While CBCT offers high-resolution anatomical details, MRI and PET can provide insights into the physiological and metabolic states of tissues.[57]

E. Challenges and Future Directions

As histopathological image analysis continues to develop and evolve, several challenges persist. With the addressing of these challenges, significant strides in integrating deep learning models, advanced imaging technologies and data analysis frameworks can be made.

1) Challenges:

a) Data Variability and Annotation: The most significant challenge in histopathological image analysis is the inherent variability in histopathological images. Variations in tissue staining, sectioning, and imaging conditions lead to significant differences in histological samples, even for the same type of tissue. This variability complicates the development of automated image analysis tools and models. Additionally, the process of annotating histological images is a labor-intensive process that requires expert knowledge, making it a difficult task to obtain high-quality annotated datasets. This limitation is especially exacerbated in the training of deep learning models which require large amounts of label data to work with satisfactory accuracy.

b) Computational Complexity: Techniques such as multi-scale feature extraction, deep learning, and multimodal image registration require significant computational resources. Processing high-resolution histopathological

images, especially whole-slide images, is computationally very expensive and time consuming. This challenge is compounded by the goal of the research, which is to integrate tools for image analysis into real-time clinical settings where rapid diagnostic decision making is needed.

c) Model Interpretability and Clinical Integration: The interpretability of deep learning models still remains a critical challenge amongst most professionals not working directly on machine learning and artificial intelligence. While these models can prove to be highly accurate in tasks such as tissue classification and feature extraction, their black-box nature makes it difficult for clinicians to understand the rationale behind specific predictions. The lack of transparency and simple understandability of the new technologies being developed is crucial for widespread and standardized application of these techniques in a clinical setting. It also remains paramount that the integration of these new technologies as a complementary tool is ensured rather than a disrupting force as these technologies need to be easy to use and integrable into standard clinical workflows.

2) Future Directions: The future of histopathological image analysis will likely focus on certain key directions such as improved synthetic data generation, optimized computational efficiency, and expanded multimodal and real-time image analysis integration. Properly annotated and standardized datasets remain scarce and it might prove to be cost-effective to develop GANs that can create realistic synthetic histopathology images to expand available training data. Additionally the usage of semi-supervised and unsupervised learning approaches will allow for the usage of large amounts of unlabeled data, reducing the dependency on manual annotations. Another important focus remains the immense computational burden the aforementioned procedures all have; the current cost to run these programs remains as a hindering barrier for entry into the field of research in resource-limited settings. The optimization of these techniques through less complex algorithms and better data storage of super high-resolution images through the usage of cloud-based platforms is essential for research in this area to thrive. Finally, the expansion of multimodal and real-time imaging combinations will play a pivotal role in the future of histopathology. The fusion of different imaging modalities such as CBCT, MRI, and PET with histological analysis will provide numerous ways to pinpoint relevant information regarding tissue pathology, leading to more accurate diagnostics. These advancements, collectively, will drive the next generation of histopathological tools.

IX. Challenges in Histopathology

Histopathology faces several challenges that impact its accuracy, efficiency, replicability, and overall effectiveness. These challenges include technical limitations, variability in practices, and the proper integration of new technologies.

A. Variability in Tissue Preparation and Staining

1) Inconsistent Staining: The process of staining tissue samples is susceptible to variability due to differences in reagent quality, timing, and technique. These inconsistencies can lead to variations in how tissues are visualized, affecting the interpretation of morphological features and potentially leading to inconsistencies and discrepancies in diagnostics.

2) Sample Quality: The quality of tissue samples can be compromised during the preparation stages which include fixation, dehydration, and embedding. These processes, when not conducted correctly, can introduce artifacts that obscure histological details and generally reduce the overall quality of the study. These artifacts complicate the diagnostic process and may result in misinterpretation or misdiagnosis.

B. Inter-observer Variability

1) Subjectivity in Interpretation: The interpretation of histopathological images can oftentimes rely on the subjective judgment of pathologists, which can lead to inter-observer variability. Different pathologists may arrive at different conclusions based on the same tissue samples, particularly in cases with complex or ambiguous lesions.

2) Standardization Issues: There are several accreditation and training guidelines set by several different institutions such as the College of American Pathologists to ensure standardized practices in laboratories and interpretations of histopathological samples but the interpretation of samples can still vary significantly between institutions and individual pathologists.

C. Technical Limitations of Imaging Techniques

1) Resolution and Contrast: Imaging techniques in histopathology still face limitations in resolution and contrast despite significant advancements in the field, particularly when analyzing thick or dense tissue sections. While super-resolution microscopy offers solutions, it is not yet widely accessible in clinical practice which leaves a gap in the ability to commonly visualize fine tissue details.

2) Depth of Penetration: Traditional light microscopy is limited in its ability to penetrate deeper tissue layers which can be overcome with techniques like multiphoton microscopy which offer deeper tissue imaging. While effective, multiphoton microscopy, with its high cost and requirement of specialized equipment, limits the widespread adoption of this technique in routine diagnostics.

D. Cost and Accessibility

1) High Cost of Advanced Techniques: Advanced histopathological techniques, such as super-resolution microscopy or multiplex immunohistochemistry, are often prohibitively expensive and not accessible to all healthcare and research settings, particularly in low and middle income countries. This cost barrier limits the adoption of new technologies.

2) Resource-Limited Settings: In many parts of the world, access to basic histopathological services is limited due to the lack of infrastructure in trained personnel, laboratory equipment, and easy access to data. These limitations can lead to delayed diagnoses and poor patient outcomes which raises questions about the need for equitable access to histopathological resources.

X. Conclusion

Histopathology continues to be critical in medical diagnostics and research with its significance bolstered by ongoing advancements in staining techniques and microscopy. This review has traced the development of histopathological dyes which have been critical in the enhancement of contrast and specificity, allowing for the detailed observation of cellular and tissue architecture essential for accurate diagnosis and research.

The incorporation of optical principles into microscopy from foundational methods to fluorescence and super-resolution microscopy has markedly improved our capability to visualize cellular structures with high fidelity. These methodologies have allowed researchers and clinicians the ability to detect subtle morphological differences that are crucial for the diagnosis of complex diseases.

Looking ahead, the integration of next-generation technologies such as virtual staining is expected to significantly elevate the resolution and depth of histopathological imaging. These innovations, coupled with the advent of computational tools in image analysis and quantification, particularly the use of machine learning algorithms, promise to enhance the precision and reliability of histopathological evaluations. Such tools can assist in automating the interpretation of complex tissue patterns, reducing observer variability and enabling more personalized and targeted approaches to treatment.

In conclusion, the progress in histopathology reinforces its central role in better understanding the complexities of disease pathology. As interdisciplinary collaboration increases, the refinement and application of these techniques will be crucial in advancing both clinical diagnostics and biomedical research- resulting in better healthcare for patients around the world. The future of histopathology is poised to make even greater contributions to medical science, improving our capabilities to diagnose, treat, and ultimately prevent diseases with unprecedented accuracy and effectiveness.

Acknowledgments

The authors would like to extend their gratitude to their respective institutions for their support throughout the development of this paper. Doruk Yalçın acknowledges Koç University, where he is an undergraduate student in the Department of Computer Engineering, and UCLA, where he is currently a Visiting Student Researcher at the Ozcan Lab. He particularly thanks the members of Ozcan Lab for providing an enriching educational environment and the opportunity to work on histopathology. Acar Karabıçak acknowledges Bilkent University, where he is pursuing his undergraduate studies in Industrial Engineering.

This work was carried out independently, without reliance on any specific lab resources, and the authors appreciate the collaborative efforts and contributions of all involved.

References

[1] M. J. Melo, "History of Natural Dyes in the Ancient Mediterranean Civilization," in *Handbook of Natural Colorants*, 2023, pp. 3–26. doi: https://doi.org/10.1002/9781119811749.ch1.

[2] A. P. Kazhdan, *The Oxford dictionary of Byzantium*. New York: Oxford University Press, 1991.

[3] H. Gest, "The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society," *Notes Rec R Soc Lond*, vol. 58, no. 2, pp. 187–201, May 2004, doi: 10.1098/rsnr.2004.0055.

[4] T. Hiramatsu *et al.*, "Iodine staining as a useful probe for distinguishing insulin amyloid polymorphs," *Sci Rep*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-73460-y.

[5] R. W. Dapson, "The history, chemistry and modes of action of carmine and related dyes," *Biotechnic and Histochemistry*, vol. 82, no. 4–5, pp. 173–187, 2007, doi: 10.1080/10520290701704188.

[6] Pasquale Chieco and Massimo Derenzini, "The Feulgen reaction 75 years on," *Histochem Cell Biol*, vol. 111, p. 345, May 1999, doi: 10.1007/s004180050367.

[7] Rosemary M Baker, "Nineteenth Century Synthetic Textile Dyes. Their History and Identification on Fabric," 2011.

[8] C. Ortiz Hidalgo, "Immunohistochemistry in Historical Perspective: Knowing the Past to Understand the Present," in *Immunohistochemistry and Immunocytochemistry: Methods and Protocols*, L. Del Valle, Ed., New York, NY: Springer US, 2022, pp. 17–31. doi: 10.1007/978-1-0716-1948-3_2.

[9] R. W. Sabnis, "Handbook Of Biological Dyes And Stains Synthesis And Industrial Applications."

[10] J. A. Titus, R. Haugland, S. O. Sharrow, and D. M. Segal, "Texas Red, A Hydrophilic, Red-Emitting Fluorophore For Use With Fluorescein In Dual Parameter Flow Microfluorometric And Fluorescence Microscopic Studies," 1982.

[11] T. C. Brelje, M. W. Wessendorf, and R. L. Sorenson, "Multicolor laser scanning confocal immunofluorescence microscopy: Practical application and limitations," *Methods Cell Biol*, vol. 20, no. 70, pp. 165–244, 2002, doi: 10.1016/s0091-679x(02)70006-x.

[12] R. B. Mujumdar, L. A. Ernst, S. R. Mujumdar, C. J. Lewis, and A. S. Waggoner, "Bioconjugate Chemistry Volume 4, Number 2 © ARTICLES Cyanine Dye Labeling Reagents: Sulfoindocyanine Succinimidyl Esters." [Online]. Available: https://pubs.acs.org/sharingguidelines

[13] S. Magaki, S. A. Hojat, B. Wei, A. So, and W. H. Yong, "An introduction to the performance of immunohistochemistry," in *Methods in Molecular*

Biology, vol. 1897, Humana Press Inc., 2019, pp. 289–298. doi: 10.1007/978-1-4939-8935-5_25.

[14] S. W. Kim, J. Roh, and C. S. Park, "Immunohistochemistry for pathologists: Protocols, pitfalls, and tips," 2016, *Seoul National University*. doi: 10.4132/jptm.2016.08.08.

[15] S. Andersson *et al.*, "Insufficient antibody validation challenges oestrogen receptor beta research," *Nat Commun*, vol. 8, Jun. 2017, doi: 10.1038/ncomms15840.

[16] M. Oron-Herman, D. Kirmayer, A. Lupp, S. Schulz, G. Kostenich, and M. Afargan, "Expression prevalence and dynamics of GPCR somatostatin receptors 2 and 3 as cancer biomarkers beyond NET: a paired immunohistochemistry approach," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-47877-0.

[17] J. Ryan, A. R. Gerhold, V. Boudreau, L. Smith, and P. S. Maddox, "Introduction to modern methods in light microscopy," in *Methods in Molecular Biology*, vol. 1563, Humana Press Inc., 2017, pp. 1–15. doi: 10.1007/978-1-4939-6810-7_1.

[18] A. D. Kim and J. B. Keller, "Light propagation in biological tissue," 2003.

[19] K. Thorn, "A quick guide to light microscopy in cell biology," *Mol Biol Cell*, vol. 27, no. 2, pp. 219–222, Jan. 2016, doi: 10.1091/mbc.E15-02-0088.

[20] M. Monici, "Cell and tissue autofluorescence research and diagnostic applications," Sep. 28, 2005. doi: 10.1016/S1387-2656(05)11007-2.

[21] J. D. Rogers, A. J. Radosevich, J. Yi, and V. Backman, "Modeling Light Scattering in Tissue as Continuous Random Media Using a Versatile Refractive Index Correlation Function NIH Public Access," *IEEE J Sel Top Quantum Electron*, vol. 20, no. 2, p. 7000514, 2013, doi: 10.1109/JSTQE.

[22] P. Haub and T. Meckel, "A model based survey of colour deconvolution in diagnostic brightfield microscopy: Error estimation and spectral consideration," *Sci Rep*, vol. 5, Jul. 2015, doi: 10.1038/srep12096.

[23] A. Sauvat, G. Cerrato, J. Humeau, M. Leduc, O. Kepp, and G. Kroemer, "High-throughput label-free detection of DNA-to-RNA transcription inhibition using brightfield microscopy and deep neural networks-NC-ND license High-throughput label-free detection of DNA-to-RNA transcription inhibition using brightfield microscopy and deep neural networks," *Comput Biol Med*, vol. 133, p. 104371, 2021, doi: 10.1016/j.compbiomed.2021.104371ï.

[24] F. Zernike, "Phase Contrast, A New Method For The Microscopic Observation Of Transparent Objects," 1942.

[25] K. Shimasaki, Y. Okemoto-Nakamura, K. Saito, M. Fukasawa, K. Katoh, and K. Hanada, "A high-resolution phase-contrast microscopy system for label-free imaging in living cells," *Cell Struct Funct*, vol. 49, no. 1, pp. 21–29, 2024, doi: 10.1247/csf.24018.

[26] E. V. Yakovlev *et al.*, "Machine learning approach for recognition and morphological analysis of isolated astrocytes in phase contrast microscopy," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-59773-2.

[27] M. Ragazzi *et al.*, "Fluorescence confocal microscopy for pathologists," *Modern Pathology*, vol. 27, no. 3, pp. 460–471, Mar. 2014, doi: 10.1038/modpathol.2013.158.

[28] M. J. Sanderson, I. Smith, I. Parker, and M. D. Bootman, "Fluorescence microscopy," *Cold Spring Harb Protoc*, vol. 2014, no. 10, pp. 1042–1065, Oct. 2014, doi: 10.1101/pdb.top071795.

[29] S. M. Hickey *et al.*, "Fluorescence microscopy—an outline of hardware, biological handling, and fluorophore considerations," Jan. 01, 2022, *MDPI*. doi: 10.3390/cells11010035.

[30] M. Izadyyazdanabadi *et al.*, "Fluorescence Image Histology Pattern Transformation Using Image Style Transfer," *Front Oncol*, vol. 9, Jun. 2019, doi: 10.3389/fonc.2019.00519.

[31] C. Elfgen *et al.*, "Comparative analysis of confocal microscopy on fresh breast core needle biopsies and conventional histology," *Diagn Pathol*, vol. 14, no. 1, Jun. 2019, doi: 10.1186/s13000-019-0835-z.

[32] J. Kang *et al.*, "Rapid tissue histology using multichannel confocal fluorescence microscopy with focus tracking," *Quant Imaging Med Surg*, vol. 8, no. 9, pp. 884–893, Oct. 2018, doi: 10.21037/qims.2018.09.18.

[33] A. D. Elliott, "Confocal Microscopy: Principles and Modern Practices," *Curr Protoc Cytom*, vol. 92, no. 1, Mar. 2020, doi: 10.1002/cpcy.68.

[34] C. S. Goldsmith and S. E. Miller, "Modern uses of electron microscopy for detection of viruses," Oct. 2009. doi: 10.1128/CMR.00027-09.

[35] H. J. Cho, H. J. Chun, E. S. Kim, and B. R. Cho, "Multiphoton microscopy: An introduction to gastroenterologists," 2011, *Baishideng Publishing Group Co*. doi: 10.3748/wjg.v17.i40.4456.

[36] E. Olson, M. J. Levene, and R. Torres, "Multiphoton microscopy with clearing for three dimensional histology of kidney biopsies," *Biomed Opt Express*,

vol. 7, no. 8, p. 3089, Aug. 2016, doi: 10.1364/boe.7.003089.

[37] E. A. Calle *et al.*, "The use of optical clearing and multiphoton microscopy for investigation of three-dimensional tissue-engineered constructs," *Tissue Eng Part C Methods*, vol. 20, no. 7, pp. 570–577, Jul. 2014, doi: 10.1089/ten.tec.2013.0538.

[38] L. Graham and J. M. Orenstein, "Processing tissue and cells for transmission electron microscopy in diagnostic pathology and research," *Nat Protoc*, vol. 2, no. 10, pp. 2439–2450, 2007, doi: 10.1038/nprot.2007.304.

[39] H. Kawasaki *et al.*, "The NanoSuit method: a novel histological approach for examining paraffin sections in a nondestructive manner by correlative light and electron microscopy," *Laboratory Investigation*, vol. 100, no. 1, pp. 161–173, Jan. 2020, doi: 10.1038/s41374-019-0309-7.

[40] L. Schermelleh *et al.*, "Super-resolution microscopy demystified," Jan. 01, 2019, *Nature Publishing Group*. doi: 10.1038/s41556-018-0251-8.

[41] S. T. Hess, T. P. K. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophys J*, vol. 91, no. 11, pp. 4258–4272, 2006, doi: 10.1529/biophysj.106.091116.

[42] J. C. M. Gebhardt *et al.*, "Single-molecule imaging of transcription factor binding to DNA in live mammalian cells," *Nat Methods*, vol. 10, no. 5, pp. 421–426, May 2013, doi: 10.1038/nmeth.2411.

[43] Eric Betzig, George H. Patterson, Rachid Sougrat, and O. Wolf Lindwasser, "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution," Sep. 2006. doi: 10.1126/science.1130258.

[44] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological Image Analysis: A Review," *IEEE Rev Biomed Eng*, vol. 2, pp. 147–171, 2009, doi: 10.1109/RBME.2009.2034865.

[45] L. Yang, P. Meer, and D. J. Foran, "Unsupervised segmentation based on robust estimation and color active contour models," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 3, pp. 475–486, Sep. 2005, doi: 10.1109/TITB.2005.847515.

[46] J. Levy, C. Haudenschild, C. Barwick, B. Christensen, and L. Vaickus, "Topological Feature Extraction and Visualization of Whole Slide Images using Graph Neural Networks the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License. HHS Public Access," 2021. [Online]. Available: https://github.com/jlevy44/WSI-GTFE

[47] S. Bhattacharjee *et al.*, "Multi-features classification of prostate carcinoma observed in histological sections: Analysis of wavelet-based texture and colour features," *Cancers (Basel)*, vol. 11, no. 12, Dec. 2019, doi: 10.3390/cancers11121937.

[48] A. Sridhar, S. Doyle, and A. Madabhushi, "Content-based image retrieval of digitized histopathology in boosted spectrally embedded spaces," *J Pathol Inform*, vol. 6, no. 1, p. 41, Jan. 2015, doi: 10.4103/2153-3539.159441.

[49] X. Li, M. Cen, J. Xu, H. Zhang, and X. S. Xu, "Improving feature extraction from histopathological images through a fine-tuning ImageNet model," *J Pathol Inform*, vol. 13, Jan. 2022, doi: 10.1016/j.jpi.2022.100115.

[50] M. M. Abdelsamea, U. Zidan, Z. Senousy, M. M. Gaber, E. Rakha, and M. Ilyas, "A survey on artificial intelligence in histopathology image analysis," Nov. 01, 2022, *John Wiley and Sons Inc*. doi: 10.1002/widm.1474.

[51] C. Chen, M. Y. Lu, D. F. K. Williamson, T. Y. Chen, A. J. Schaumberg, and F. Mahmood, "Fast and scalable search of whole-slide images via self-supervised deep learning," *Nat Biomed Eng*, vol. 6, no. 12, pp. 1420–1434, Dec. 2022, doi: 10.1038/s41551-022-00929-8.

[52] D. Tellez *et al.*, "Whole-Slide Mitosis Detection in H&E Breast Histology Using PHH3 as a Reference to Train Distilled Stain-Invariant Convolutional Networks," *IEEE Trans Med Imaging*, vol. 37, no. 9, pp. 2126–2136, Sep. 2018, doi: 10.1109/TMI.2018.2820199.

[53] L. Jose, S. Liu, C. Russo, A. Nadort, and A. Di Ieva, "Generative adversarial networks in digital pathology and histopathological image processing: A review," Jan. 01, 2021, *Wolters Kluwer Medknow Publications*. doi: 10.4103/jpi.jpi_103_20.

[54] M. Boktor, J. E. D. Tweel, B. R. Ecclestone, J. A. Ye, P. Fieguth, and P. Haji Reza, "Multi-channel feature extraction for virtual histological staining of photon absorption remote sensing images," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-52588-1.

[55] N. Pillar, Y. Li, Y. Zhang, and A. Ozcan, "Virtual Staining of Nonfixed Tissue Histology," May 01, 2024, *Elsevier B.V.* doi: 10.1016/j.modpat.2024.100444.

[56] T. Puentes-Morelos, V. Simancas-Escorcia, A. Tarón-Dunoyer, C. M. Ardila, and A. Díaz-Caballero, "Histological Evaluation of Internal Dental Resorption: An Analysis of a Cohort of 50 Cases," *Int J Dent*, vol. 2024, 2024, doi: 10.1155/2024/1454079.

[57] R. S. Vanguri *et al.*, "Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer," *Nat Cancer*, vol. 3, no. 10, pp. 1151–1164, Oct. 2022, doi: 10.1038/s43018-022-00416-8.

RESEARCH

KUUJE
Koç University Undergraduate Journal of Engineering

Anemica: A Mobile Application for Non-Invasive Detection of Anemia Using Palm Images

Berfan Pirhan
Department of Computer Engineering
Faculty of Engineering
Koç University
Istanbul, Turkiye
bpirhan20@ku.edu.tr

Esra Duran
Department of Computer Engineering
Faculty of Engineering
Koç University
Istanbul, Turkiye
eduran20@ku.edu.tr

Rawnaq el Kindy
Department of Computer Engineering
Faculty of Engineering
Koç University
Istanbul, Turkiye
rkindy@ku.edu.tr

Remziye Sude Güngör
Department of Computer Engineering
Faculty of Engineering
Koç University
Istanbul, Turkiye
rgungor19@ku.edu.tr

***Abstract*— Anemia is a disease that impacts the lives of individuals on a daily basis, consisting of many types and various symptoms including fatigue, shortness of breath, and weakness. For the detection of anemia, the standard practices have been mostly invasive for patients, which include obtaining blood samples. To simplify this procedure, non-invasive procedures have been developed, which mainly consist of the analyses of images of patients' palms, fingernail beds, or conjunctival pallor. In this paper, we propose a mobile application solution for anemia detection, Anemica, which analyzes the palm images of patients using a trained convolutional neural network (CNN) model and returns anemia estimates and hemoglobin ranges based on the results. This is accomplished by analyzing 16 regions of the palm image, including finger landmarks, with each region categorized into one of four groups based on the calculated erythema indices. The erythema index serves as an indicator of the redness of the respective regions. Our solution integrates findings from the literature review and demonstrates substantial potential for enhanced accuracy and marketability.**

Keywords*— *Adam Optimizer, Anemia, Complete Blood Count (CBC), Convolutional Neural Network (CNN), Decision Trees, Erythema Index, F1 Score, Hemoglobin (hB), Image Augmentation, Image Classification, k-Nearest Neighbor (kNN), Naïve-Bayes (NB), ReLU Activation, RGB, Softmax Activation, Support Vector Machine (SVM)

I. Introduction

Anemia is a condition which affects the lives of many individuals ranging among all ages and genders. There are many causes of anemia, such as iron deficiency, vitamin deficiency, inflammation-based anemia, aplasticanemia, bone marrow disease-related anemia, hemolytic anemia, and sickle cell anemia [1]. Furthermore, the symptoms of anemia can include shortness of breath, fatigue, pale or yellowish skin, irregular heartbeat, chest pain, dizziness, and headaches [1]. These symptoms and variations of anemia make the condition an indispensable matter for health. Anemia has a considerable impact on children and pregnant women [2]. Based on this prevalence and significance of anemia-related conditions, it becomes necessary to detect anemia and treat it as soon as possible.

Anemia is usually detected among patients undergoing invasive methods, such as drawing blood samples to obtain parameters including the complete blood count (CBC) and the shape/size of red blood cells [3]. Further information may also be requested, such as the medications used by patients as well as their weight [3]. This traditional approach can cause difficulties for patients who suffer from anemia, which has sparked research on non-invasive detection methods

Non-invasive detection of anemia comprises of detecting the anemia status of a patient without the need to engage in physical procedures like drawing blood from patients. After conducting a literature review, we have evaluated numerous research papers concerning non-invasive solutions. These papers analyze body parts such as palm images [4], [5], fingernail beds [9], conjunctival pallor [7], and their combined analysis in children [10]. All of them have shown promising results ranging from 92% to 99% detection accuracy; however, the opportunity of creating a full-fledged product for end-users stemming from such research has not yet been properly explored.

In this paper, we demonstrate our project Anemica, which combines the methodology from existing research for non-invasive detection of anemia-related conditions into a mobile application with high prospects for anemia detection and marketability. It uses a trained convolutional neural network (CNN) model to assign 16 landmarks on a palm image to 4 severity categories based on the erythema indices for each, which indicate the redness of a region. Based on the results, it provides anemia estimates and hemoglobin ranges for each severity category obtained, through a fluid and user-friendly interface. Our approach can be further enhanced with more enriched datasets and training, as it currently faces issues with detection accuracy due to dataset limitations.

II. Related Works

There exists a plethora of research investigating non-invasive methods of anemia detection. Studies in the literature have focused on image-based approaches for the detection of anemia, in which various body parts are analyzed. These include the analysis of palm images [4], [5], conjunctival pallor [7], fingernail beds [9], and their combined analysis in children for iron-deficiency anemia [10], and they have shown promising results regarding detection accuracy and performance.

Manuscript Received: 09.06.2024 | Manuscript Reviewed: 09.08.2024 | DOI: 10.5281/zenodo.13375850

A. *EfficientNet CNN Models with Palm Images*

Amrutesh et al. proposed a convolutional neural network (CNN) based approach to categorize hemoglobin levels of patients using their palm images [4]. This approach establishes a machine learning-based solution, where the images are classified as either being anemic or non-anemic. The EfficientNetB0 model with softmax activation and Adam optimizer demonstrated the best results, with a 97.52% accuracy and an F1 score of 97%. Additionally, their paper has tested the entire EfficientNet model family, ranging from B0 to B7, with a consistent accuracy of over 92% achieved for all. However, the same models using linear activation have reported to be underperforming. As per the data obtained, their study showed a promising approach for anemia detection with high detection accuracy.

B. *Ensemble Models with Palpable Palm Images*

Appiahene et al. utilized palpable palm images obtained from patients in their study [5]. The dataset used for this study consisted of 710 participants' images obtained from various hospitals in Ghana [6]. These images were extracted and converted to RGB percentiles for training. A hybrid model was developed using ensemble learning models, where the stacking ensemble model obtained the highest accuracy of 99.92%, among testing of other ensemble models. Their study utilized various algorithms including Naïve-Bayes (NB), decision trees, k-nearest neighbor(kNN), support vector machine (SVM), and random forest (RF). The random forest approach gave the highest individual accuracy of 99.53%and an F1 score of 99.61%. Moreover, the study combined the random forest model with the Naïve-Bayes algorithm to get an accuracy rate of 99.93% and a 100% rate on all other parameters including the F1 score. Despite the Naïve-Bayes algorithm resulting with less accurate results where the accuracy remained at 78.16%, the combination of the NB and RF models resulted with a near- perfect accuracy for the classification of palm images. Their study demonstrated an accurate approach for classifying anemic and non-anemic palm images.

C. *Conjunctival Pallor Analysis with Phone Cameras*

Collings et al. focused on the conjunctival pallor images of patients to determine their anemia status [7]. Their study obtained images of patients' palpebral conjunctivae, which is a mucous membrane that covers the inner surface of the lower and upper lids of the eyes [8]. Phone cameras were used in ambient light, and the hemoglobin concentration was obtained by analyzing the pictures using the erythema index of the conjunctiva, which indicates the redness of the region. The conjunctival erythema indices had a sensitivity of 93% for the training set and 57% for the validation set, with specificity values for detecting anemia obtained as 78% and 83%, respectively.

D. *Patient-Sourced Fingernail Bed Image Analysis*

Mannino et al. approached the non-invasive anemia detection problem through analyzing the fingernail bed images provided by patients [9]. Their study made use of a phone camera for analysis, and solely utilized patient-sourced photos through a mobile application that analyzes hemoglobin levels of fingernail bed photos. The application detected anemia, where the hemoglobin levels were less than 12.5 g/dL. The accuracy this study obtained was ±2.4 g/dL, with a sensitivity rate of 97% when compared with complete blood count (CBC) hemoglobin levels. Furthermore, the application achieved a better accuracy of ±0.92 g/dL when personalized calibration was utilized.

E. *Comparative Multiregional Anemia Detection*

Asare et al. came up with a machine learning-based approach to determine whether the analysis of conjunctival pallor, palpable palms, or fingernail images provides a more accurate result for detecting anemia in children [10]. They focused on detecting iron-deficiency anemia and applied the Naïve Bayes, SVM, CNN, kNN, and decision tree approaches. The dataset shares the source with [6], and was collected, preprocessed, and then the detection model was developed. The dataset sizes they obtained were equal for all three regions, which also apply to the augmented version. They have achieved promising results, with the CNN model yielding the best outcome with 98.45% accuracy for the conjunctiva, 98.33% accuracy for the fingernails, and 99.12% accuracy for the palpable palms. The F1 scores, precision levels, and recall levels for the CNN model have also been the most promising, followed by the Naïve Bayes, decision tree, kNN, and SVM algorithms.

III. Methodology

A. *Project Background and Aims*

With Anemica, we set out to develop a mobile application that would be user-friendly and would detect the anemia status of a patient along with hemoglobin estimates based on the anemia status obtained. Despite the existence of relevant literature for non-invasive detection, there currently is an undiscovered opportunity to simplify the anemia detection procedure for patients through the integration of relevant findings within a mobile application designed to be user-friendly and with appealing visuals. We aimed for the application to work on both iOS and Android phones and designed the application in a cross- platform setting for this purpose. Our application currently holds high potential for future improvements and deliverability for end-users. Its design is momentarily focused on the use of patients; however, the app will also contain an interface for doctors, and a structure connecting patients and doctors will be established upon other features depending on the future potential and prospects for our application.

B. *Data Collection and Availability*

For the data collection procedures of our project, we have faced various limitations regarding the availability of relevant data for our project design. Parts of the structural basis for our design and most of the dataset we were supplied were provided courtesy of Koç University Hospital, thanks to the support of our advisors Prof. Çiğdem Gündüz Demir and Prof. Öznur Özkasap of the Computer Engineering Department of Koç University, as well as Prof. Özlem Yalçın of Koç University School of Medicine. The dataset we have used for our project were provided with their permission.

To further discuss on data availability, it can be stated that there are available options for finding datasets for palm images, which includes the contents of the ensemble model research [5]. The dataset utilized for [5] was obtained from a publicly available dataset containing 4260 images in total and is available for downloading [6]. However, since the dataset is from hospitals in Ghana, it was not suitable for our purposes as our approach also utilizes the RGB values obtained, and we could have obtained skewed results for patients with lighter skin colors. As a result, we consulted the Koç University Hospital to obtain palm image samples of anemic and non-anemic patients.

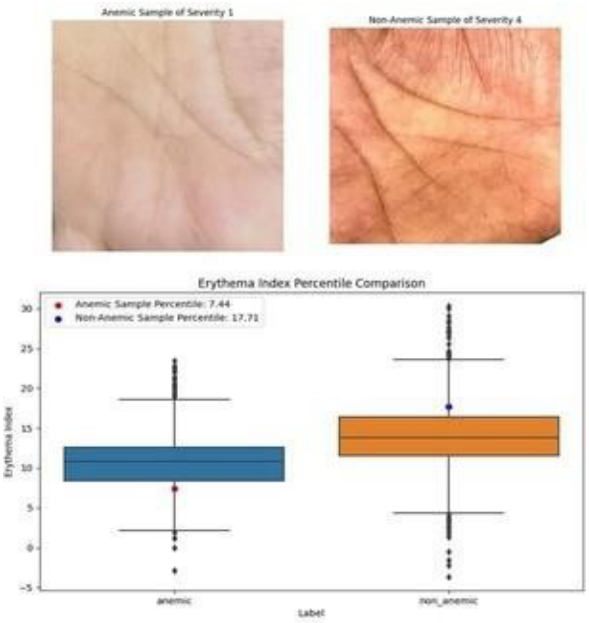


Fig. 1. Anemic and non-anemic palm image samples with theirerythema index percentiles highlighted.

This dataset was categorized into four, depending on the severity of the anemia for each, which we have also used in our project as our main approach. After filtering through images with unsuitable lighting conditions, image quality, and distance, we were left with 51 images for training. Sample images with severity values 1 and 4 with their erythema index distribution percentiles can be observed in Fig 1. The top left image belongs to an anemic patient, and the top right image belongs to a non-anemic patient. The graphs show the distribution of erythema indices of anemic patients with severity categories 1 and 2 on the left, and non-anemic patients with severity categories 3 and 4 on the right, where the distributions are based on an augmented version of the dataset. The boxes represent 50% of the data in their respective groups. The average erythema index of anemic samples is lower compared to the non-anemic samples. The percentiles of the sample anemic and non-anemic images are highlighted with red and blue dots, respectively. Outlier values for the calculated erythema indices are also present at the ends of the graphs.

Including the augmented version, this remaining dataset of 51 images contained images of the same hand in different angles. When combining these limitations, we have faced certain drawbacks with obtaining accurate results for our project. However, these can be mitigated with enriched datasets and more specialized training approaches. Possible improvements for the study will be discussed in the upcoming sections.

C. Project Design

We have designed our application Anemica as part of three main components, including the frontend system, backend system, and the image processing system. The application currently supports the patient interface where after creating their accounts, patients are able to scan their palms using their own cameras or by selecting an image from their gallery. We also have initiated a doctor interface which is currently in a primitive state. With the doctor interface, we seek to implement features where doctors are able to create new patients in the patient database. We also strive to initiate specialized scanning features where either the doctor or the patient can select specified regions of their hand and scan for selected regions.

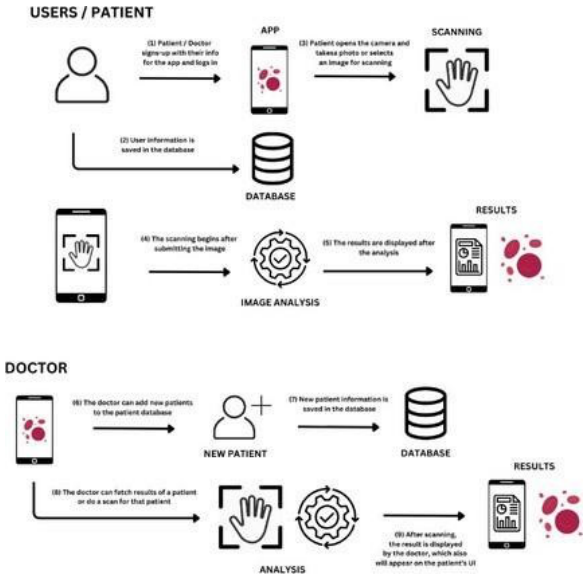


Fig. 2. The general flow and interface plan of Anemica.

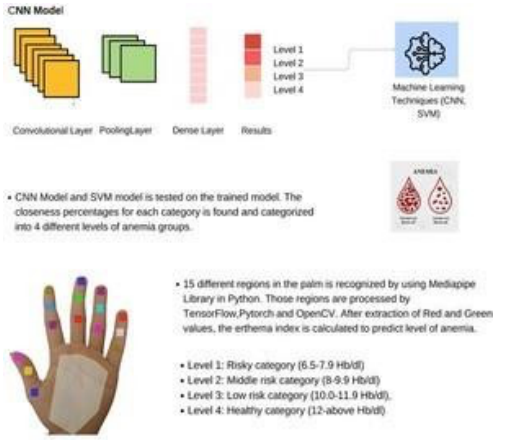
The current status of our application incorporates most of the structure plans we have established. The overall diagram demonstrating the planned application can be found in Fig. 2. Patients or general users are able to register to the application, scan their hands and get results on their anemia status, hemoglobin estimates, and recommendations based on the results. While not being fully implemented, doctors are currently able to access the results of an existing patient in the patient database. The separation of doctors and patients and the addition of the remaining planned features for the doctor interface are planned for future implementations.

In this paper, we will be focusing on the details and procedures of the image processing part. We have used Python as the main programming language for the image analysis. For the detection of landmarks on the palm, we have used the MediaPipe library offered by Google [11]. We utilized the

Hand Landmarker task for MediaPipe, which has hand landmark detection for still images as one of its features.

Fig. 3. Each landmark on the left hand provided by Media Pipe HandLandmarker [11].

The library detects 21 landmarks on the hand, with 4 on each finger and one for the base of the hand. Each region detected by the Hand Landmarker is displayed in Fig. 3 [11]. To achieve our goal, we reduced the total regions to 16 to accommodate for the palm region better by combining the landmarks 0, 1, 5, 9, 13, and 17 into one polygonal area. As for the rest of the landmarks, we define rectangular regions with the center being the landmark detected for each of those regions. Concerning the case of augmenting the data we have obtained, we utilized TensorFlow, which we conducted due to the existing limitations of the dataset at hand. After these steps, we used the OpenCV library to create and highlight the



aforementioned rectangular and polygonal regions and extract the RGB values for each region.

Fig. 4. The layer diagram of our CNN model and anemia result categorization.

The CNN model we have used classifies a given image into one of the four categories from which the model was trained on. The detailed structure of this model is described in Fig. 4. We initially tested the use of a support vector machine (SVM) classifier to categorize the images based on the RGB parameters but did not obtain promising results. As a result, we utilized a CNN model instead. The accuracy results acquired from the CNN model and a comparison between SVM and CNN models can be seen as displayed in the results section.

Our CNN model makes use of a convolutional layer, a pooling layer, and a dense layer, which then outputs the category for the image given. The categories are numbered from 1 to 4, where 1 indicates the highest severity and 4 indicates the lowest. The model maps the three dimensions of RGB into 1 using the parameters of ReLU activation for the hidden layer and the softmax activation for the output layer. In order to obtain the best classification, our model used the Adam optimizer. This categorization procedure is applied for each of the 16 regions on the palm and achieved through the calculation of the erythema index for each region.

As described in Equation 1, the erythema index E in our system is obtained from the division of the logarithm value of the red channel intensity Ired and the green channel intensity Igreen. This formula was based on the erythema index calculation used to determine the absorption of red and green light by the hemoglobin in scar tissue and healthy skin [12]. However, it is important to note that the conjunctival erythema index calculations in [7] differ slightly from our formula, where the logarithm values of the red and green channels were calculated separately. The accuracy of our model can also be influenced by this difference, which can be validated with further testing.

$$E = 100\, x\, log(\frac{Ired}{Igreen}) \qquad (1)$$

The overall output of our model is an array that contains the percentage for each region based on the erythema index, pertaining to the category assigned, along with the category number for each region. We then export this array through our backend system to the frontend to map the categories for each region in order to display anemia probabilities and hemoglobin ranges.

```
1. FUNCTION calculate_results(indices, loaded_model)
2.    SET results to array of 32 zeros
3.
4.    FOR i FROM 0 TO LENGTH(indices) - 1
5.       IF indices[i] is not NaN THEN
6.
7.          SET X to [[indices[i]]]
8.          SET df to DataFrame(X, columns=['E'])
9.
10.         SET r to loaded_model.predict(df[['E']], verbose=0)
11.
12.          SET results[i] to 100 * max(r) / sum(r) # Calculate
accuracy
13.         SET results[i + 16] to index of max(r) # Record the
category
14.      END IF
15.   END FOR
16.
17.   RETURN results
18. END FUNCTION
```

Algorithm 1. Calculation function for the image processing system

In the image processing system, the main calculation is carried out within the result calculation function, which is given in Fig. 5. We take the erythema indices and the trained model as parameters, where we load the model from a .h5 file. We use NumPy and Pandas to predict the category for the erythema index for each region. Then, we load the model and find the accuracy percentages as the first 16 elements of the final array, and the category number obtained in the remaining

part of the array. The hemoglobin ranges we return are assigned to each category, where the riskiest category 1 returns a range of 6.5-7.9 g/dL, 2 returns 8.0-9.9 g/dL, 3 returns 10.0-11.9 g/dL, and 4 returns 12+ g/dL.

IV. Results

During the course of our project design, we managed to achieve promising results, although there is significant room for improvement. As per considerable dataset limitations as well as the lack of thorough exploration of other approaches for training due to time constraints, we managed to obtain strained accuracy levels for our model.

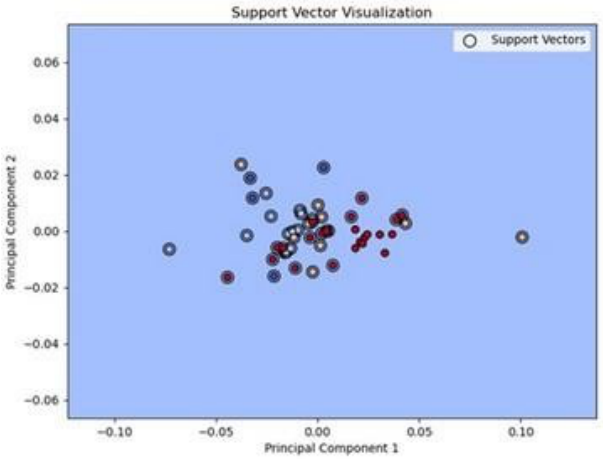


Fig. 5. The support vector visualization of RGB parameters for ourdataset.

Before training the CNN model, the SVM classifier we used to classify our images obtained a 45% accuracy for the original image set and a 40% accuracy for the augmented set during training. The SVM visualization for the RGB parameters of our dataset can be found in Fig. 6. Consequently, we opted for the CNN model which maps the three RGB dimensions into one using ReLU and softmax activation functions, with the Adam optimizer to obtain better classification results.

This model obtained a 48% accuracy for the training phase, and a 46% accuracy for the testing phase, where relevant regions to corresponding severity categories were assigned. For the CNN approach, we divided the dataset into training and testing sets with a 20% test ratio. We tested our model on the dataset divided into four severity categories. However, our model was not very successful on assigning the images to each of the severity categories, where the output of our tests mainly consisted of predicting the severities of images as 2 or 4. To improve the precision of our model's classification, we have tried different approaches to our training logic, as well as trying out different augmentation approaches. Despite these attempts, the precision of our model has mostly remained the same, especially with category predictions. Due to this outcome, as well as to obtain clearer metrics for our model, we have decided to measure the performance our model with a binary approach. In this approach, we checked if the category predictions of 1 or 2 were matching with anemic images, and if the category predictions of 3 or 4 were matching with non-anemic images. Using this approach, we have obtained an overall detection accuracy of 70.6%, a recall rate of 41.8%, a precision rate of 63.1%, and an F1 score of 50.3% for our CNN model.

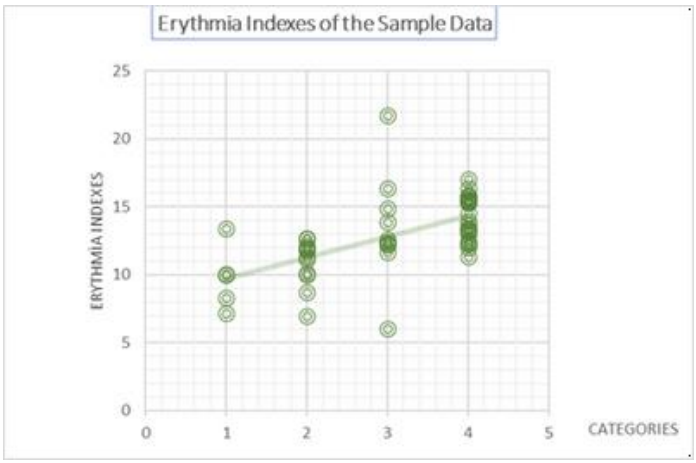


Fig. 6. The correlation chart of the erythema indices of the original dataset for each severity category.

The erythema indices we obtained for each image for our original dataset also show a correlation, with some outliers present. The chart that demonstrates this distribution is given in Fig. 7, where we observe a correlation for the erythema index obtained for each image based on the category they belong. As the categories progress from 1 to 4, the overall erythema indices obtained are increasing. This occurs as the overall redness of the images increase with the reduction of the anemia risk. This chart demonstrates the consistency for our approach using erythema indices of the palms to detect anemia and hemoglobin values.

In addition to our approach for the detection of anemia and hemoglobin values, we have also achieved promising results for the interface and usability of our application. Example screenshots for our application, including the result pages which are connected with our image processing system can be found in the Appendix. Alongside providing a comfortable way for detecting anemia, we have also managed to create a dynamic, appealing, and user-friendly interface to enhance the prospects for our project Anemica. Once a user creates their account and logs in, they are greeted with the main dashboard of our application. In this dashboard, users can see the result of their latest test, informative cards about the causes of anemia and how it can be reduced, and widgets mentioning how to scan and previous results which take the user to the analysis page and the result page, respectively. When the analysis page is launched, the user is greeted with a tour on how to get ready for a scan. Once the tour is completed, the user is prompted with two options consisting of selecting an image from their gallery or taking a new photo. Once an image is selected or a photo is taken, the user can send that photo for analysis.

The image analysis currently takes around 30 seconds to 1 minute to fully complete. When the analysis page is launched, the user is greeted with a tour on how to get ready for a scan. Once the tour is completed, the user is prompted with two options consisting of selecting an image from their gallery or taking a new photo. Once an image is selected or a photo is taken, the user can send that photo for analysis. The image analysis currently takes around 30

seconds to 1 minute to fully complete processing the images and displaying the results, along with the regional highlighted version of the image provided in the results page. Users will be able to see their results in the results page once the analysis is completed, where they can see additional information upon tapping on an existing result. In addition to these pages, we also have a profile page in which users can see their information including their name, age, and gender, with a profile photo section which is currently not functional. From the profile page, users are currently only able to log out of the application and delete their accounts

V. Discussion

A. Comparison

The results obtained within existing literature have mostly provided highly accurate results with the help of the combination of different machine learning algorithms and training approaches. Our model has observed a significantly strained training accuracy level of 48% compared to existing studies, which have mainly achieved training accuracy levels of over 90%.

Our model has suffered from other metrics as well, with a recall rate of 41.8%, a precision rate of 63.1%, and an F1 score of 50.3%, in addition to a more moderate detection accuracy of 70.6%. However, it is important to note that we worked with a very limited dataset consisting of 51 images. With the provision of considerably higher numbers of images or datasets, when accounted for adequate lighting and distance conditions for palm images of anemic and non- anemic patients, we can achieve much better results for our project and model.

B. Limitations and Improvements

As discussed previously, our CNN approach and current anemia detection system faces various limitations. Besides the data limitations, we also faced issues with obtaining images with appropriate lighting conditions and angles. The training set we used to train our model only consisted of palm images, whereas we also provided results based on finger landmarks for a more comprehensive analysis. We used the model trained based on palm images for the finger landmarks, which has also impacted the accuracy of our approach. To account for these shortcomings, our model can be refined with an approach where each of the 16 regions are trained separately, with the help of an enriched dataset. Using pretrained models can also be a viable option for our case, if we are unable to resolve dataset issues. We have also utilized the detection of palm images instead of other regions like the conjunctivae or fingernail beds for the purpose of simplifying the detection procedure for patients and doctors. Also seen in previous pages, there exists a multitude of studies focusing on analyzing palm images for non-invasive anemia detection, which has also been an important parameter in our decision for Anemica.

In addition to improvements for our image processing system, we also plan improvements for the existing functionalities of our application on top of improvements for the user interface, depending on the future prospects of our project. For instance, the full integration of the doctor interface would be a monumental addition to our application to enhance the patient-doctor interactions on detecting their anemia conditions and treating them in a more comfortable fashion. Feature improvements for privacy concerns including the ability to delete all data for an existing patient as well as the separation of doctor and patient information are also among our reinforced plans to ensure an enhanced product with promising utility.

VI. Conclusion

Anemia is a condition which impacts the lives of many individuals regardless of age or gender. In the traditional approach, patients are diagnosed through invasive means that involve drawing blood, which can cause distress for the patients. Consequently, non-invasive detection approaches of anemia conditions have gained monumental importance. With our project Anemica, we combined existing literature of non-invasive detection methodology to come up with a marketable and comfortable anemia detection application using palm images. To provide a broader perspective, we utilized erythema index calculations for finger landmarks in addition to the palm. Despite existing limitations for Anemica, there is significant room for improvement for accuracy. As part of identifying anemia, our application Anemica demonstrates high potential and prospects for becoming a valuable tool for patients and healthcare professionals.

Acknowledgments

Over the course of our project, we had a valuable experience working closely with our advisors Prof. Çiğdem Gündüz Demir and Prof. Öznur Özkasap from the Computer Engineering Department of Koç University, and Prof. Özlem Yalçın from Koç University School of Medicine. We would like to express our gratitude and appreciation to them for providing their support and assistance throughout the development of this project.

REFERENCES

[1] Mayo Clinic, "Anemia - Symptoms and Causes," *Mayo Clinic*, May 11 2023.https://www.mayoclinic.org/diseases-conditions/anemia/symptoms-causes/syc-20351360

[2] World Health Organization, "Anaemia in women and children," *www.who.int*, 2021. https://www.who.int/data/gho/data/themes/topics/anaemia_in_women_and_children

[3] Mayo Clinic, "Anemia - Diagnosis and treatment - Mayo Clinic," *Mayoclinic.org*, 2023. https://www.mayoclinic.org/diseases-conditions/anemia/diagnosis-treatment/drc-20351366

[4] A. A, A. Amrutesh, G. B. C G, A. R. K P, and G. S, "EfficientNet models for detection of anemia disorder using palm images," *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1437–1443, Apr. 2023. doi:10.1109/icoei56765.2023.10125656

[5] P. Appiahene *et al.*, "Application of ensemble models approach in anemia detection using images of the palpable palm," *Medicine in Novel Technology and Devices*, vol. 20, p. 100269, May 2023. doi:10.1016/j.medntd.2023.100269

[6] J. W. Asare, P. Appiahene, and E. Donkoh, "Anemia Detection using Palpable Palm Image Datasets from Ghana," *data.mendeley.com*, vol. 1, Jun. 2022, doi: https://doi.org/10.17632/ccr8cm22vz.1

[7] S. Collings *et al.*, "Non-invasive detection of anaemia using digital photographs of the conjunctiva," *PLOS ONE*, vol. 11, no. 4, Apr. 2016. doi:10.1371/journal.pone.0153286

[8] "Conjunctiva - an overview | ScienceDirect Topics," *www.sciencedirect.com*. https://www.sciencedirect.com/topics/medicine-and-dentistry/conjunctiva

[9] R. G. Mannino *et al.*, "Smartphone app for non-invasive detection of anemia using only patient-sourced photos," *Nature Communications*, vol. 9, no. 1, Dec. 2018, doi: https://doi.org/10.1038/s41467-018-07262-2

[10] Justice Williams Asare, P. Appiahene, Emmanuel Timmy Donkoh, and G. Dimauro, "Iron deficiency anemia detection using machine learning models: A comparative study of fingernails, palm and conjunctiva of the eye images," May 2023, doi: https://doi.org/10.1002/eng2.12667

[11] "Hand landmarks detection guide | Edge," *Google for Developers*. https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker

[12] M. E. H. Jaspers and P. Moortgat, "Objective Assessment Tools: Physical Parameters in Scar Assessment," *PubMed*, 2020. https://www.ncbi.nlm.nih.gov/books/NBK586073

KUUJE
HIGHLIGHTS

KUUJE
Koç University Undergraduate Journal of Engineering

AI: Is It a Protector From Unfairness or Just a Blanketed Bias Machine?

Elif Ezerel[1]
Department of Psychology
Faculty of Social Sciences and Humanities
Koç University
Istanbul, Turkiye
eezerel20@ku.edu.tr

***Abstract*— As artificial intelligence (AI) becomes integrated into an increasing number of aspects of our lives, it is clear that AI has the power to revolutionize decision-making processes. This study focuses on how skewed data sets and algorithms could promote ableism and prejudice against disabled applicants. Additionally, it discusses the use of facial recognition technology in hiring procedures, the lack of oversight, and the need to consider the variety of disabilities. Finally, the conclusion addresses the limitations of AI in representing the diversity and fluidity of human identity and experiences.**

Keywords*— *Artificial intelligence, Algorithmic bias, Ableism, Technoableism, Disability discrimination, Bias in AI, Ethical AI

Artificial intelligence (AI) has the ability to completely transform decision-making processes, and this is evident in a time when AI is being incorporated into more and more facets of our lives. AI advocates claim that by reducing or even eliminating human biases, their system may produce more equal and objective results [1]. This optimistic perspective, however, ignores a crucial problem; the ingrained biases in the algorithms themselves. These prejudices frequently support discrimination and inequality, especially towards people with disabilities. The hiring process is one instance where this issue is most noticeable. Even though AI is supposed to be unbiased, the data and training algorithms that are employed might unintentionally reinforce preexisting biases, which can be detrimental to job applicants with disabilities. This paper focuses on how these biased data sets, and algorithms, may enable ableism and discriminate against disabled people during recruitment processes.

Ableism is the explicit or implicit preference for non-disabled bodies and brains, as well as the negative attitudes that follow towards disability and individuals with disabilities. Our expectations and views are shaped by ableism. These biases against disability and handicapped individuals are evident in the physical environment as well as in the institutional and societal norms that influence day-to-day existence [2].

The behaviours and circumstances that marginalise or prevent the contributions of disabled people, remove them from communal life, and reject first-hand accounts of their experiences are all rooted in ableism. One such example is technoableism. The term "technoableism" refers to a discourse on disability that both emphasises the use of technology to empower disabled people while promoting ableist stereotypes about who is deserving and what kinds of bodies and minds are desirable. Artificial Intelligence and machine learning are the most recent wave of technological remedies to the perceived issues faced by individuals with disabilities. AIs that are being developed now have several instances of ableism [2].

Many definitions of artificial intelligence base their meaning on analogies with human thinking. According to these definitions, AI is sophisticated technology intended to mimic human intellect. AI is also capable of carrying out activities that previously needed human involvement. In reality, AI is just software with the ability to learn, decide, complete tasks, and solve problems [3]. Moreover, as it has been said to be almost ‘bias free’, most companies are starting to use it in their recruitment programs [4]. The increasing prevalence of automated employment solutions can be attributed to advancements in machine learning and data analytics. By incorporating AI into the recruiting process—from locating applicants and evaluating applications to automating interviews—employers want to maximise hiring efficiency. Hiring AIs compare applicants to previous workers or look for personality factors linked to high performance to find the ones who are most likely to "fit" or flourish in the job. Endorsers assert that their algorithms enable firms to increase recruiting diversity since they are not only more effective than human recruiters but also less biassed. Despite the excitement surrounding AI's ability to boost diversity in hiring, many vendors rely on ambiguously defined notions of bias and inadequately represented data, which make it difficult to understand how bias might be reflected in AI and potentially worsened [5].

Manuscript Received: 17.07.2024 | Reviewed: 08.09.2024 | DOI: 10.5281/zenodo.13375891

AI uses a set of data, also known as its training data, to teach the AI how to learn, decide and solve problems. Algorithms, which are essentially a collection of instructions that can be followed step-by-step to arrive at a certain output, are necessary for AI to function. Algorithms, like all technology, should be viewed as manifestations of power and control and as a system rather than as discrete objects. Even while algorithms are frequently promoted as a means of measuring fairness and making decision-making more objective, there are many different interpretations of what constitutes "fairness," many of which are mutually exclusive or entail substantial trade-offs [6].

AI programmes may have different effects when trained on nonrepresentative datasets, even in cases where material specifically referencing protected classes is not included. It is true that datasets may exaggerate the associations between belonging to a group and unfavourable outcomes. Vulnerable groups are more likely to have their mistakes observed, documented, and utilised as training data for algorithms since they are frequently the targets of excessive scrutiny. For instance, because information on handicapped persons is disproportionately obtained in settings related to addiction, homelessness, and violence, language processing AI links the mention of disability to those conditions. In addition, algorithms may not be able to identify patterns in datasets that underrepresent particular groups; instead, they may concentrate on the majority in order to prevent overfitting. However, the methods of issue framing, proxy selection, and data labeling are also responsible for algorithmic bias, in addition to unrepresentative datasets and erroneous input–output linkages. Making vague requirements, like "fit" and "employability," comprehensible to AI is a process shaped by societal conventions and attitudes rather than being impartial. For instance, if AI systems that simulate "employability" in terms of good posture, fluent communication, and consistent eye contact were flawless, they would discriminate against a large number of candidates with disabilities. When presented with scant or heterogeneous data, data scientists must also make arbitrary judgements about the selection of proxy variables or the aggregation of attributes. This might lead to introduction of biases, loss of accuracy and poor policy and decision-making. [5]

It has been stated that the recruitment system will undoubtedly be fair if it is unaware of a person's handicap status. Employers may believe that by outsourcing the evaluation to an ostensibly “neutral” AI and doing away with traditional human prejudice, they are assisting candidates with disabilities. That being said, this ignores the likelihood that white, able-bodied men invented the instruments in the first place [7]. The issue lies in the fact that the impairment frequently affects other data points that are given in the model. A blind person will need more time to browse the page before they can respond to the inquiry if the test programme is poorly constructed [8].

If that time isn't taken into account when evaluating them, then everyone with a comparable impairment using the same instrument will be systematically disadvantaged. This is a result of norm learning in machine learning systems. They don't give outliers any particular treatment; instead, they optimise for norms. However, those with impairments don't always fit the mould [8].

Another example is recruitment videos. Among these businesses offering video interviews is HireVue, which creates technologies utilised by almost one hundred well-known corporations, like Unilever, JP Morgan, and Shell. By rating candidates based on a variety of criteria, the programme employs artificial intelligence (AI) to assist in the analysis of job video interviews and identify the most qualified applicant [9]. As a consequence, psychological judgements about a person's likelihood of succeeding in a position are drawn from their facial data. A person with a facial difference is likely to score lower when facial recognition technology is employed during the interview process [10]. This will result in more discrimination against disabled people. There are people with facial disfigurement due to their diseases, people who had strokes cannot often control one side of their faces or blind people may not even look at the camera during their interviews. AI may be fair in its judgement, however literature and field results show that the dataset that AI is trained upon is biased.

Ableism is a topic people claim to know about but in reality we haven’t even scratched the surface. This topic was interesting for me on a personal note because I am also disabled and use a wheelchair. I have experienced being the other in our society firsthand, especially in recruitment. However, this is about to be even more exacerbated by the use of AI. It’s a valiant effort to minimize and in the future hopefully eliminate all the bias disabled people experience in recruitment with the use of AI tolls. This doesn’t seem possible at the moment. However, people should also realise the fault isn’t only in the process but also in the datasets that companies train their AI tools on. We cannot just say we have done our due duty of making AI fair and leave a developing part of technology alone to develop with these preexisting biases. Incorporating disability perspectives into AI discussions reveals AI's limitations in capturing the fluidity and complexity of human identity and experiences. With this discussion, hopefully even more companies will be inspired to discuss these topics and show the preexisting and possible problems.

REFERENCES

[1] Seppälä, P., & Małecka, M. (2024). AI and discriminative decisions in recruitment: Challenging the core assumptions. *Big Data & Society, 11*(1), 20539517241235872.

[2] Shew, A. (2020). "Ableism, Technoableism, and Future AI." *IEEE Technology and Society Magazine, 39*(1), 40–85.

[3] Ohlheiser, A. W. (2023, June 5). *AI automated discrimination. Here's how to spot it.* Vox. https://www.vox.com/technology/23738987/racism-ai-automated-bias-discrimination-algorithm

[4] Lee, A. (2019, November 26). *An AI to stop hiring bias could be bad news for disabled people.* WIRED UK. https://www.wired.com/story/ai-hiring-bias-disabled-people/.

[5] Tilmes, N. (2022). Disability, fairness, and algorithmic bias in AI recruitment. *Ethics and Information Technology, 24*(2), 21.

[6] Moura, I. (2023). Encoding normative ethics: On algorithmic bias and disability. *First Monday.*

[7] Alexiou, G. (2020, December 13). *Algorithmic and AI assessment tools are a new frontier in disability discrimination.* Forbes. https://www.forbes.com/sites/gusalexiou/2020/12/13/algorithmic-and-ai-assessment-tools---a-new-frontier-in-disability-discrimination/?sh=13da69f544f7

[8] Hao, K. (2018, November 28). *Can you make an AI that isn't ableist?*. MIT Technology Review. https://www.technologyreview.com/2018/11/28/1797/can-you-make-an-ai-that-isnt-ableist/

[9] Lee, A. (2019, November 26). *An AI to stop hiring bias could be bad news for disabled people.* WIRED UK. https://www.wired.com/story/ai-hiring-bias-disabled-people/.

[10] Humans for AI. (2023, February 19). *Ableism in AI: Facial Recognition Technologies in Recruitment Processes, the Lack of Governance and Account for the Multiplicity of Disability.* Medium. https://medium.com/@humansforai/ableism-in-ai-facial-recognition-technologies-in-recruitment-processes-the-lack-of-governance-67c36a1a0745

[11] "Hand landmarks detection guide | Edge," *Google for Developers*. https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker

KUUJE
Koç University Undergraduate Journal of Engineering

Legacy in Hematology and Cancer Immunology: An Interview with Prof. Dr. Emin Kansu

Doruk Yalçın
Department of Computer Engineering
Faculty of Engineering
Koç University
Istanbul, Turkiye
dyalcin21@ku.edu.tr

Çağan Aras Incel
TED Ankara College
Ankara, Turkiye

***Abstract*—This interview with Prof. Dr. Emin Kansu explores his remarkable career in hematology and cancer immunology, highlighting his role in the establishment of the Hacettepe Cancer Institute. Prof. Dr. Kansu reflects on his academic journey from Hacettepe University to the United States, where he specialized in blood diseases and immunology, and his return to Turkey to contribute to the country's medical advancements. The discussion includes the challenges of academic mentorship, the evolution of medical research in Turkey, and the future prospects of fields such as bioinformatics, genomic medicine, and stem cell research. Prof. Dr. Kansu also shares his insights on the ethical considerations surrounding the use of artificial intelligence in medicine, highlighting the need for ongoing dialogue and regulation in this rapidly advancing field.**

Keywords—Hematology, Cancer Immunology, Hacettepe University, Medical ethics, Medicine in Turkey

I. Introduction

Prof. Dr. Emin Kansu's contributions to the fields of hematology and cancer immunology have left an indelible mark on medical research and education in Turkey. A graduate of Hacettepe University Faculty of Medicine, Prof. Dr. Kansu pursued advanced training in the United States, where he specialized in internal medicine, hematology, and immunology. Upon his return to Turkey, he played a critical role in establishing the Hacettepe University Cancer Institute, a pioneering institution that has become a cornerstone for cancer treatment and research in the country.

Throughout his career, Prof. Dr. Kansu has been deeply committed to academic mentorship, guiding countless students and young researchers. His efforts have significantly contributed to the development of medical education and research infrastructure in Turkey. In this interview, Prof. Dr. Kansu shares his insights on the evolution of medical research, the challenges and opportunities in academic mentorship, and the ethical considerations surrounding emerging technologies like artificial intelligence in medicine.

As a recipient of prestigious awards, including the TÜBİTAK Science Award, Prof. Dr. Kansu's reflections provide valuable perspectives on the current state and future directions of medical research in Turkey and beyond. This conversation offers an opportunity to gain a deeper understanding of the experiences and vision of one of Turkey's leading medical researchers.

II. Could You Introduce Yourself?

I graduated from Hacettepe University Faculty of Medicine in 1970. Shortly after, I went to the United States for specialization training and residency. I spent a year working as an intern there and completed my residency. After completing my post-graduate education in the U.S., I returned to Hacettepe University in 1980. During my time in the U.S., I completed training in internal medicine, followed by further training in hematology and immunology.

Upon my return to Hacettepe, I began advancing in academia—first becoming an associate professor and later a full professor. It was during this period that we conceived the idea of establishing a Cancer Institute at Hacettepe University, an institution that would focus on both the prevention, treatment, and research of cancer, a significant need for both the university and Turkey at the time. Alongside my colleagues, we successfully founded the Hacettepe University Cancer Institute, which operates as a major institution on campus. With Professors Dinçer Fırat, Namık Çevik, and Şevket Ruacan I also played a key role in its foundation and the training of its faculty members. We established clinical departments within the institute, including medical and pediatric oncology, preventive radiation oncology, nuclear medicine and tumor pathology. By 2003, the Cancer Institute became fully operational with its inpatient and outpatient services.

Throughout this process, we received tremendous support from many people, notably the late Mr. Süleyman Demirel President of the Turkish Republic at the time, who appreciated the necessity of this Institute and provided crucial assistance. The university administration was also highly supportive, and many of my students who were trained in the Cancer Institute have now become professors, specializing in various fields. Many of them spent two or three years in the U.S. before returning, and today, they are leading authorities in their respective areas.

I was particularly involved in the establishment of cancer research and basic oncology and even spent two years abroad to help set up the bone marrow transplant unit. Today, we can

Interview Conducted: 16.08.2024 | Interview Reviewed: 26.08.2024 | DOI: 10.5281/zenodo.13376244

Prof. Dr. Emin Kansu

proudly say that the Hacettepe University Cancer Institute serves as a reference center in Turkey for both pediatric and medical oncology. This has been a significant part of my academic life, and I am very happy with what we've achieved.

I retired in 2014 due to my age and looking back, I can say that we have made significant contributions to medicine, hematology, oncology, pathology, and academics in Turkey. The Institute now receives around 4,000 new cancer patients every year, providing diagnoses and treatment plans for both children and adults and offering an invaluable service. Every unit in the country consults with this Institute and refers patients here for diagnosis and treatment. During my time there, my work was purely academic; I didn't open a private practice. I focused on education, research, and patient care within the Cancer Institute.

In state universities, the mandatory retirement age is 67, while in private universities like Koç, you can work beyond 70, as far as I know. Although my days now are less intense than before after retiring, I'm staying academically active by visiting national universities, giving lectures, at Jefferson Medical College, reviewing projects, and evaluating award applications.

III. Your work in hematology and cancer immunology is quite remarkable. What specific topic in these fields excites you the most and keeps you engaged in your research?

I enjoyed every aspect of my work. As academics, we all specialize in a particular field. We can't afford to be scattered, saying, "I'll look into this, and I'll look into that," especially in medicine. You need to receive training and become an expert in a single area. It's similar to engineering, where you must specialize in a specific field; it's impossible to know everything. I was particularly interested in blood diseases and the immune system problems that arise from them. These two issues are crucial mechanisms in the development of cancer.

You might ask, for example, "Why didn't you focus on brain tumors?" That could have been an option, of course. However, since my education in the U.S. was focused on blood diseases, I naturally entered this field. But it's tough to pick just one area among the ones I've worked on, especially when you enjoy what you do. My research primarily focused on blood production. I explored the problems that occur in blood production from the bone marrow, how blood is produced normally, and how it is produced under disease conditions. In medical terms, we call this phenomenon 'Hematopoiesis'. This is still not fully understood; we haven't completely unraveled the mechanisms of its control but we're working on it and have made significant progress.

For example, one of my distinguished world-famous mentors in the U.S. discovered a hormone that influences this process—a hormone called Erythropoietin, which is secreted by the kidneys. It travels to the bone marrow and controls the production of red blood cells. I was very much interested in this process and spent four or five years studying it. I researched the mechanisms by which this hormone affects the bone marrow, how it initiates blood production, and its biological properties, and I published several papers on this topic.

Back then immunology was just beginning to gain recognition—this was around 1978-79. This new field was generating a lot of excitement, and alongside blood production, I became very interested in it as well. After all, immune cells are produced during the blood production process. This was a very enjoyable research journey for me, and when I returned to Turkey, I assisted students with their theses and research projects on this subject.

IV. While you were a student at Hacettepe University Faculty of Medicine, was there a specific event or moment that directed you toward such a successful academic career? During this period, did you have a mentor or role model who inspired you?

It's impossible to pinpoint a single event; anyone who claims otherwise is just telling a tale. In university education, it's a group of professors who guide you; you can't just pick one because they all complement each other's teachings. We were fortunate at Hacettepe because the university was founded in 1964, and we were its first students. As the first students, we were led by Professor İhsan Doğramacı, who founded the Hacettepe University. Foreseeing the establishment of this faculty, he had sent many young academics to the United States in the mid-1950s. When these

individuals returned, they formed the academic staff of our university, even before becoming associate professors. They were working with great enthusiasm, having just returned, the faculty was brand new, the course topics were freshly developed, and the students were new as well, so they provided us with an excellent education. Our class size was also small, just 64 students. This was an ideal number at the time, allowing the professors to pay close attention to each of us, get to know us well, identify any deficiencies that might arise, and address them, ultimately giving us an outstanding medical education over six years.

Of course, as students, we observed our professors—what they were teaching, which areas they emphasized, and how they conducted their lectures. Since they had returned from America, they educated us according to the American medical education model of 1968. Because of these conditions, it's tough to single out one person or event. As we entered our final year, in 1969-70, during our last year as interns, working like doctors before officially becoming doctors, our professors encouraged us to continue our education in the United States. Not everyone went, of course, but 20-25 of us did, having seen how our professors had been trained there and how deeply it had impacted them. We took the American exams of that time, like the ECFMG, and scored high marks. We graduated in 1970 and went to the U.S. for further education. However, the foundation of our academic education was still laid at Hacettepe.

Of course, not all graduates from our term became academics; some worked in hospitals, and others opened private practices. Not everyone is required to be an academic; it's a profession that requires dedication. It requires a commitment to student education, teaching, and feeling responsible for a student's learning, which is why I don't intend to criticize those who don't pursue academia. However, many of our classmates who graduated with us continued as academics. In this regard, our professors at Hacettepe provided mentorship for us, and by the 1970s, Hacettepe had started to establish itself as a model for other universities in the country. However, I should mention Professors Şeref Zileli and Faruk Özer as my distinguished mentors who convinced me to receive my post-graduate training in internal medicine and hematology.

V. After graduating, you continued your education and research at Thomas Jefferson University. What was it like for you to work abroad at a young age? What differences did you observe between the medical field in Turkey and abroad? What are your thoughts on the advantages and disadvantages of these differences?

That's a difficult question. People might say “If America excels in science, why can’t we?”. Well, we can’t. How many gold medals did the U.S. win in the Olympics? Forty, the same as China. We can't compare ourselves to a country like that. Their athletes have been training for years, extensively preparing physically before every competition. It's a country with a population of 336 million, built on a very disciplined and education-oriented culture that’s built on decades of hard work and progress, and apart from China, right now, no other country has that. China, of course, isn't as open to the world, but it’s still able to compete with the U.S. I won’t compare Turkey to America; that would be a mistaken endeavor. But I’d rather tell you why I decided to come back to Turkey. I experienced America, understood all its aspects during my 12 years there, and then decided to return. When I decided to return my boss and my colleagues said, “Don’t go back,” but I insisted saying “No, I’m going back. I won’t stay in America.” I love the US, but I rejected becoming American. I understand and appreciate the United States, but I felt that I was obligated to bring back what I learned back to Turkey and to the Turkish students. When my professor asked me why I was returning, I said, “I come from a middle-class, civil-servant family. And even though my family wasn’t wealthy they still managed to have me educated at TED Ankara College, which was expensive for families even back then. They sacrificed from themselves, even cut on their meals, but they still managed to have me and my sibling educated. When I was accepted to Hacettepe Medical School and studied there for six years I truly realized how much Turkey had invested in my education. I could stay in the US, but why would I? An American student has lots of resources, he can find 300-500 professors like me, but a Turkish student can’t. If I can provide this education to Turkish students and give back to them what I had received, I could be of more use.” My department head respected that greatly.

Hacettepe University Cancer Institute, 2006

Students often ask me, "Should I go to America and stay there? “ I can’t answer that question. You should go to America; as a computer engineer, you should see the work they’re doing and stay there for as long you want. However, the decision, whether to stay or to return, is a personal one that no professor can make this decision for you. You’re at UCLA right now and might be thinking “I’ve learned everything I set off to learn and got what I needed from here, I’m going back to Turkey.” That’s a very respectable choice but no one can say anything against you either if you choose to stay. We personally chose not to stay. But we still attend conferences, and teach periodically at U.S. universities, and that's enough for us. Seeing a student listen to you attentively, focused on what you're passionately explaining, is a wonderful thing. If I or your professors at Koç University had wanted to stay in America, we could have done so. Your professors at Koç are highly valued; I know them and the internal structure of the university well become I worked there. The medical school is excellent with a faculty comprising of very prominent names

and they didn't stay. Ultimately they chose to educate the students at Koç University and make a meaningful impact on their lives because that was what made them happy with their contributions and satisfied with their careers.

VI. How did you feel when you won the TÜBİTAK (The Scientific and Technological Research Council of Türkiye) Science Award? How did this award impact your scientific work?

We obviously don't work for the sake of winning awards; I've never worked with the goal of receiving an award. Some people may be more eager about it, but that's not really the right approach to have. The work you do might be noticed and appreciated by some professors and faculty members, and they may decide to give you an award. Naturally, you feel happy when your work is recognized; it's satisfying to know that your efforts are appreciated, and it can be demoralizing if they aren't. If I see someone doing good work but not being recognized, I make sure to bring it to the attention of the relevant people, who then evaluate the person's work independently. But as I said, we don't work for awards. The TÜBİTAK Science Award is the highest scientific award in Turkey; winning it gives us an extra drive, a motivation to ask ourselves, "Can I do even more research, be even better?" Awards provide motivation; people who receive them continue their work with renewed energy. I received the TÜBİTAK Encouragement Award in 1979 and the TÜBİTAK Science Award in 1997. At that time, our professors evaluated our work and found it deserving; but, even if I hadn't won the award, I would have continued working.

VII. You've served under the umbrella of TÜBA (The Turkish Academy of Sciences) for many years. Based on your experiences during this time, what kind of structural changes and infrastructure improvements do you believe should be made to more effectively support and advance scientific research in Turkey? In your opinion, which areas should be prioritized?

I'm the first full member of TÜBA, and currently, I'm an honorary member. I still visit universities across Turkey and this year, I'll be visiting places like Bitlis University, Süleyman Demirel University, and 9 Eylül University. In my visits I meet with rectors, deans, and students, trying to understand how they are doing. For example, I visited Kahramanmaraş University, and you have to understand that there's a significant difference between the conditions there and 9 Eylül University. The professors at universities like Kahramanmaraş University only work there for a while and some of them may move on to other universities While I intend to visit more universities in Eastern and Southeastern Anatolia, I see that the biggest issue in these universities is the lack of mentors and experienced professors. How can science progress without mentors- people who guide and educate you? At Koç University you have an abundance of people who are there for you but the general and most significant problem in several universities is the lack of qualified faculty, which is a problem that is even more critical than financial issues. You need someone to transfer their

TUBITAK Science Awards

knowledge to you and us at the Turkish Academy of Sciences we are aware of these issues and can see it clearly.

In Turkey, we have both well-established universities like Istanbul University and new ones like Bitlis University. Turkey currently has **208 universities and 81** medical faculties. This is a significant topic of debate in terms of education policy that we're not going to go into now, but simply put, you can't support these many universities because there aren't enough professors. There may be new academics, but they aren't fully qualified yet. For me, this number is too high. We can't train enough professors to support these many universities. This issue is also a financial one. Let's say a student says, "I'm satisfied with my education, but I also want to do research." They then go to the dean for financial support who tells them that they simply don't have the budget for the research. At that point, to get budgeting, you need to go to TÜBITAK and secure a project with them. This is a path that ultimately limits a researcher and their project proposals might still rejected. I know of an assistant who, four or five years ago, sold his car to finance his thesis; there is simply not enough money in the system to finance a big research proposal. When you establish so many universities, you have to create the necessary infrastructure needed for the employment of faculty staff, the teaching staff, and the budget that they need. This is beyond us as individuals but it's part of a broader education policy. Koç University is a luxurious university both for students and staff but if you were to step out of that environment and go to another university, say for example Niğde University, you would be surprised by the differences and challenges. It's just that the people there work in a different environment with more challenges and hurdles.

Doruk: As you've mentioned, I've seen this disparity while working at UCLA as well. The people who conduct research here are simply brilliant and are very successfully working at the very top of their fields globally. I don't want to insinuate that their only advantage is monetary, but even the prices of the computers they use and the costs that these labs can sustain are incredibly high. A majority of the universities in Turkey simply can't begin seriously working in certain fields because of the cost of entry.

Yes, you're right. Would you say that the costs are too far out of reach for us? Can't we provide these opportunities in Turkey as well?

Doruk: We can of course, if we were to focus on certain fields and areas but with so many universities with numerous laboratories and projects… it seems almost impossible to establish conditions in every lab that are at a satisfactory base level right now. Maybe if we were to focus on a number of labs and fund them extensively, these conditions could be met.

What you're saying is very important. In previous years, I suggested, and we discussed, taking TÜBİTAK's science center in Gebze as a model for academic research centers in Turkey. Instead of each university incurring expenses and sustaining niche operations individually, I proposed establishing research centers in each of the geographic regions of Turkey. It didn't happen, for reasons I'm not sure why. We might not be able to provide the same conditions as some American universities across the board, but by establishing scientific centers in places like Bursa or Eskişehir, we can create centralized state-of-the-art facilities accessible to all universities. This is something that can be done.

TÜBITAK: Marmara Research Center

VIII. What developments do you foresee in the fields of hematology/immunology in the near future? Are there any projects you're following that excite you?

You're somewhat involved in this topic already. I can mention a few key areas in our field that excite me, though it's hard to rank them. Bioinformatics, which is relevant to both medicine and engineering, is one of them. Biotechnology, specifically the discovery of drugs and the identification and study of molecules that determine our health and diseases, is a rapidly evolving field. In 1953, the DNA in our cells was discovered, and it was awarded the Nobel Prize in 1962. Then, in 2001, the Human Genome Project mapped the human genome from a collaboration between the U.S. and U.K. scientists. These two developments significantly advanced both medicine and engineering, marking a turning point for us. After 2001, a field called Genomic Medicine emerged, which has been advancing in many areas, from drug development to pathobiology and disease treatments.

The stem cell, which is the most fundamental cell, that forms every part of the human body was discovered. These cells, found in utero form the entire structure of the baby and later the human. Now, we can produce these stem cells in laboratories, which has emerged as a new field. It is called induced pluripotent stem cells (iPS Cells). For example, there are high-level studies seeking answers to how organs like the liver, intestines, and brain are formed. Stem cell research has also started and is in good progress at Hacettepe. Alongside this, 3D printing has emerged as a new field, and there is ongoing work on artificial organ production. It hasn't been fully realized yet, but these developments may happen in the next 10-15 years.

There's also the field of neurology, which has long been unresolved, such as with Alzheimer's disease. This topic remains an unsolved mystery. However, important studies on this are being conducted at Koç and Bilkent Universities, as well. I believe there will be significant breakthroughs in this area over the next 15 to 20 years. These fields interest me greatly and are advancing rapidly. I saved cancer for last because, despite many advancements, we still haven't solved it. Cancer will continue to occupy us for a long time; it's a very complex disease. We know certain factors that increase risk, like smoking and UV light but there are other unknown factors as well. Cancer can affect everyone, from infants to adults, making it very difficult to map the entirety of it. Then there's the field of bioinformatics, which engineers are currently working on extensively. And more recently, there's artificial intelligence, a field that has somewhat surpassed my era. I don't want to talk too much about it because I don't fully understand it, but it has very important ethical dimensions. For example, is it ethical to write a research paper using ChatGPT? Is it a fraud? The future of this is quite challenging. Is it ethically correct for doctors to use artificial intelligence? This is being debated. Its use impacts everyone—the patient, the insurance company, the health care provider, and the hospital.

Doruk: There are certain tools developed that aid in the decision-making process of diagnosing diseases. If something were to go wrong, a doctor using AI could rightfully say "I didn't make the decision, the black box did." Who is responsible if an error occurs, the developers, the doctors? Or how will we resolve the legal and ethical problems that arise when we're trying to collect data to train and perfect our models?

I think when you return to Turkey, you should organize a symposium on this topic. Invite professors, faculty members, students, and professionals. It's important that we put these issues out on the table and debate them with experts from different fields and have a rational discussion. We currently need a collective brainstorming session because we're at the very start of the process and we have many problems that need to be solved, or at least reach a consensus on. Thank you for inviting me to this interview and for your kind support and encouragement during the session