DAFTAR ISI

Halaman Judul ……..……..... i
Halaman Pengesahan ……... ii
Identitas Dan Uraian Umum ..…… iii
DAFTAR ISI..... 1
RINGKASAN PROPOSAL..... 2
BAB 1. PENDAHULUAN 3
BAB 2. TINJAUAN PUSTAKA 5
BAB 3. METODOLOGI PENELITIAN 12
BAB 4. JADWAL PENELITIAN..... 20
BAB 5. RENCANA ANGGARAN BIAYA (RAB)..... 21
DAFTAR PUSTAKA 22
LAMPIRAN 1. ROAD MAP PENELITIAN 26
LAMPIRAN 2. SURAT KESEDIAAN BERKOMITMEN 27

RINGKASAN PROPOSAL

Judul: Benchmarking Model Vision-Language Komersial dalam Klasifikasi Kanker Payudara pada Citra Mikroskopi Histopatologi. **Permasalahan**: Perkembangan model vision-language (VLM) komersial dalam beberapa tahun terakhir memungkinkan analisis citra medis tanpa pelatihan khusus pada domain tertentu. Model-model ini mampu menerima input gambar dan menghasilkan interpretasi tekstual secara langsung (zero-shot). Namun, kemampuan intrinsik model tersebut dalam mengenali kanker pada citra histopatologi, khususnya pada mikroskopi kanker payudara, belum banyak dievaluasi secara sistematis. Sebelum model semacam ini dipertimbangkan sebagai alat bantu diagnostik, diperlukan evaluasi kinerja yang objektif dibandingkan dengan standar referensi klinis, yaitu ahli patologi. **Tujuan**: Penelitian ini bertujuan mengevaluasi dan membandingkan performa tiga model vision-language (VLM) komersial yaitu Anthropy Claude Sonnet 4.5, GPT Vision 4o, dan Gemini ProVision dalam mengklasifikasikan gambaran histopatologi kanker payudara menggunakan pendekatan zero-shot benchmarking, dan diagnosis ahli patolog sebagai gold standard. **Metode**: Desain cross-sectional retrospektif dilakukan pada dua kohort yaitu dataset publik dan dataset internal serta dua tingkat magnifikasi menggunakan pembesaran 100x dan 400x untuk menilai stabilitas performa antar skala dan antar sumber data. Setiap gambar diklasifikasikan secara biner (benign/malignant) dan multi-kelas oleh masing-masing VLM menggunakan prompt terstandarisasi. Performa dinilai berdasarkan akurasi, sensitivitas, spesifisitas, presicion, F1-score, dan Cohen's kappa. **Kesesuaian dengan road map**: penelitian ini sesuai dengan dengan road map penelitian ketua peneliti dan Road Map Penelitian Fakultas. **Mitra**: pengambilan data juga diambil dari bagian Patologi Anatomik RSUD Prof. Dr. Soekandar Mojosari, Mojokerto, Jawa Timur. **Luaran** yg direncanakan: berupa poster terkait penelitian serta sertifikat HKI, artikel ilmiah yang terpublikasi pada jurnal nasional terakreditasi.)

Kata Kunci :
B*enchmarking*, histopatologi, kanker payudara, kecerdasan artifisial, *Vision-Language model*

BAB 1. PENDAHULUAN

1.1. Latar Belakang

Kanker payudara masih menjadi keganasan dengan insidensi tertinggi pada perempuan secara global. Menurut World Health Organization (Singh et al., 2023), kanker payudara merupakan kanker paling sering didiagnosis pada perempuan di dunia dan menjadi penyebab kematian yang signifikan setiap tahunnya. Diagnosis definitif kanker payudara tetap bergantung pada pemeriksaan histopatologi melalui evaluasi morfologi jaringan, terutama dengan pewarnaan Hematoxylin-Eosin (H&E).

Interpretasi sediaan histopatologi bersifat subjektif dan dapat dipengaruhi oleh pengalaman serta beban kerja patolog. Variabilitas antar-pemeriksa masih dilaporkan terjadi, khususnya pada lesi borderline atau dalam penentuan grading tumor. Seiring berkembangnya digital pathology dan whole slide imaging (WSI), volume data citra meningkat secara drastis, sehingga membuka peluang sekaligus tantangan dalam pemanfaatan teknologi komputasi. Pemanfaatan Artificial Intelligence (AI), khususnya deep learning berbasis convolutional neural networks (CNN), telah menunjukkan performa yang menjanjikan dalam analisis citra histopatologi kanker payudara. Studi oleh Spanhol et al. (2016) menunjukkan bahwa pendekatan deep learning mampu mengklasifikasikan citra histopatologi kanker payudara (benign vs malignant) dengan akurasi yang tinggi menggunakan dataset BreakHis (Spanhol et al., 2016).

Selain itu, Coudray et al. (2018) menunjukkan bahwa model deep learning tidak hanya mampu mengidentifikasi tipe kanker dari slide histologi, tetapi juga dapat memprediksi karakteristik molekuler tertentu langsung dari citra jaringan (Coudray et al., 2018). Pendekatan multiple instance learning (MIL) juga banyak digunakan dalam analisis WSI karena mampu menangani citra berukuran sangat besar dengan membagi slide menjadi patch kecil, lalu mengagregasikan informasi tersebut untuk prediksi akhir. Campanella et al. (2019) melaporkan bahwa sistem deep learning berbasis MIL pada whole slide image dapat mencapai performa setara ahli patologi dalam deteksi kanker pada berbagai organ, termasuk payudara (Campanella et al., 2019).

Perkembangan terbaru mengarah pada model Vision-Language (VL), yang mengintegrasikan representasi visual dan teks dalam satu kerangka multimodal. Model seperti CLIP (Contrastive Language–Image Pretraining) diperkenalkan oleh Radford et al. (2021) dan menunjukkan bahwa representasi bersama gambar dan teks dapat meningkatkan generalisasi model pada berbagai tugas klasifikasi tanpa pelatihan spesifik domain (Radford et al., 2021). Nguyen et al (2024) memanfaatkan retrieval teks dari VLM untuk analisis kuantitatif dan interpretabilitas. Performa klasifikasi dan klastering sebanding dengan model visual konvensional, yang berarti menunjukkan adanya potensi explainable zero-shot pathology (Nguyen et al., 2024).

Dalam konteks medis, pendekatan serupa mulai diterapkan untuk menghubungkan citra dengan laporan klinis, seperti pada model multimodal radiologi yang menggabungkan gambar dan teks laporan untuk meningkatkan pemahaman kontekstual (Zhang et al., 2020).

Dalam bidang histopatologi, pendekatan Vision-Language berpotensi menghubungkan fitur morfologi jaringan dengan deskripsi laporan patologi, seperti grading, status reseptor hormon, atau subtype molekuler. Pendekatan ini tidak hanya berpotensi meningkatkan akurasi klasifikasi, tetapi juga dapat meningkatkan interpretabilitas model karena hasil prediksi dapat dikaitkan dengan konsep tekstual yang lebih mudah dipahami oleh klinisi. Meskipun demikian, implementasi model Vision-Language dalam histopatologi kanker payudara masih memerlukan validasi lebih lanjut, terutama terkait generalisasi model, kebutuhan dataset pasangan gambar–teks yang terkurasi, serta aspek etika dan regulasi dalam penggunaan AI pada praktik klinis.

Sebagian besar penelitian kecerdasan buatan dalam histopatologi berfokus pada model deep learning yang dilatih khusus (task-specific CNN). Sebaliknya, evaluasi terhadap model komersial tertutup yang tidak dilatih ulang pada dataset spesifik masih terbatas. Sebelum model semacam ini

dipertimbangkan sebagai alat bantu diagnostik, diperlukan evaluasi kinerja yang objektif dibandingkan dengan standar referensi klinis, yaitu ahli patologi.

Penelitian ini bertujuan mengevaluasi dan membandingkan performa tiga model vision-language komersial dalam klasifikasi kanker payudara pada citra mikroskopi histopatologi menggunakan pendekatan zero-shot benchmarking. Desain dilakukan pada dua kohort (dataset publik dan dataset internal) serta dua tingkat magnifikasi untuk menilai stabilitas performa antar skala dan antar sumber data.

1.2. Perumusan Masalah

Dari paparan diatas maka dapat dijabarkan rumusan masalah sebagai berikut:

1. Bagaimana perbandingan performa tiga model vision-language komersial yaitu Anthropy Claude Sonnet 4.5, GPT-4 Vision, dan Gemini ProVision dalam mengklasifikasikan kanker payudara pada citra mikroskopi histopatologi?
2. Bagaimana kesesuaian (Cohen's kappa) model dengan gold standar ahli patologi?
3. Apakah performa masing-masing model konsisten pada dua tingkat magnifikasi (100x dan 400x)?
4. Apakah performa model stabil pada dua kohort data (Dataset publik dan dataset internal) yang berbeda?

4.1. Tujuan Penelitian

Tujuan Umum

Mengevaluasi dan membandingkan kemampuan tiga model vision-language komersial yaitu Anthropy Claude Sonnet 4.5, GPT-4 Vision, dan Gemini ProVision dalam klasifikasi kanker payudara menggunakan pendekatan zero-shot.

Tujuan Khusus

1. Menghitung akurasi, sensitivitas, spesifisitas, precision, dan F1-score masing-masing model.
2. Mengukur kesesuaian (Cohen's kappa) antara hasil model dan gold standard ahli patologi.
3. Membandingkan performa model pada magnifikasi 100× dan 400×.
4. Membandingkan performa model pada dataset publik dan dataset internal.

4.2. Batasan Penelitian

Batasan penelitian ini Adalah:

1. Slide tumor payudara ganas berupa Karsinoma invasive-NST, sedangkan lesi jinak berupa Fibroadenoma mammae.
2. Gambar/Citra yang digunakan adalah foto mikroskopi dan bukan *scanning/Whole Slide Imaging*.
3. Tiga model vision-language komersial yang digunakan yaitu Anthropy Claude Sonnet 4.5, GPT-4 Vision, dan Gemini Pro Vision.

4.3. Kaitan Penelitian dengan Road Map Penelitian Pribadi dan Road Map Penelitian Fakultas

Penelitian pribadi ini selaras/sejalan dengan roadmap penelitian fakultas yaitu bidang kedokteran dengan menerapkan pemanfaatan ilmu biomedik, termasuk Teknik biomedik didalamnya termasuk pemanfaatan kecerdasan artifisial, serta pengaplikasian ilmu epidemiologi klinik khususnya untuk kasus onkologi, yang sejalan dengan road map penelitian pribadi.

BAB 2. TINJAUAN PUSTAKA

Epidemiologi Tumor Payudara

Kanker payudara masih menjadi keganasan dengan insidensi tertinggi pada perempuan secara global. Laporan terbaru dari *Global Cancer Observatory* menunjukkan bahwa kanker payudara telah melampaui kanker paru sebagai kanker paling sering didiagnosis di dunia yaitu mencapai 2,3 juta kasus baru per tahun dan menjadi penyebab utama kematian akibat kanker pada perempuan (Sung et al., 2021). Di Indonesia maupun secara global, beban penyakit ini tidak hanya diukur dari angka insidensi, tetapi juga dari implikasi jangka panjang terhadap kualitas hidup dan kebutuhan terapi sistemik.
Insidensi kanker payudara terus meningkat secara global, terutama di negara dengan transisi ekonomi dan gaya hidup modern (Arnold et al., 2022). Angka kelangsungan hidup berbeda antar negara, dengan survival lebih tinggi di negara berpendapatan tinggi dibandingkan negara berkembang karena perbedaan skrining dan terapi(Allemani et al., 2018).

Kanker payudara merupakan kanker terbanyak pada perempuan di Indonesia, dengan estimasi lebih dari 65.000 kasus baru per tahun (Sung et al., 2021) Sebagian besar pasien di Indonesia datang dalam stadium lanjut, yang berkontribusi terhadap tingginya angka mortalitas (Gautama, 2022; Pasaribu et al., 2020). Faktor risiko yang meningkat di Indonesia meliputi perubahan gaya hidup, obesitas, dan penurunan angka kelahiran, sejalan dengan tren global (Arnold et al., 2022).

Tumor payudara jinak jauh lebih sering terjadi dibandingkan kanker, terutama pada wanita usia muda, dan kebanyakan benjolan payudara pada wanita usia dibawah 40 tahun merupakan lesi jinak (Zeinomar et al., 2019). Fibroadenoma adalah tumor jinak paling umum, terutama pada perempuan usia <35 tahun (Pavithra D et al., 2023). Lesi jinak tertentu seperti proliferative lesions dengan atypia (atypical ductal hyperplasia dan atypical lobular hyperplasia) meningkatkan risiko kanker payudara di masa depan secara signifikan sekitar 3–5 kali lipat (Hartmann et al., 2005).

Fibroadenoma merupakan diagnosis tersering pada benjolan payudara di klinik Indonesia, terutama pada usia remaja dan dewasa muda yaitu pada usia kurang dari 40 tahun (Khomeini, 2024; Utamie et al., 2023). Mayoritas lesi payudara yang diperiksa secara histopatologi di fasilitas pelayanan kesehatan Indonesia adalah jinak, meskipun proporsi keganasan meningkat pada usia di atas 40 tahun (Hartmann et al., 2005; Khomeini, 2024).

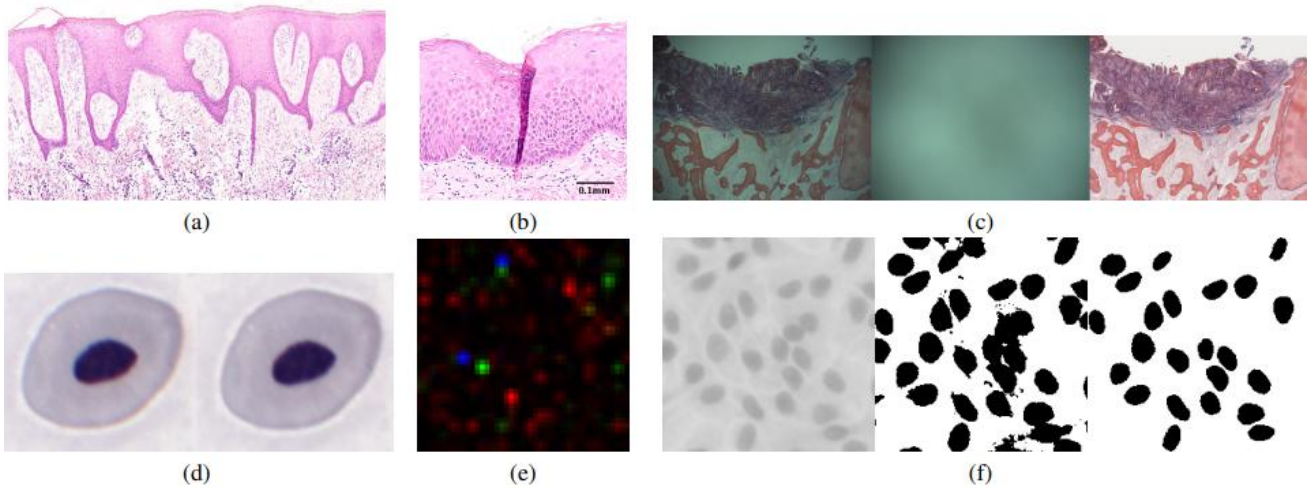
Kanker Payudara dan Peran Histopatologi

Terlepas dari kemajuan imaging dan biomarker molekuler, diagnosis definitif kanker payudara tetap bertumpu pada evaluasi histopatologi jaringan yang diwarnai Hematoxylin–Eosin (H&E) (Allison et al., 2020). Pada preparat mikroskopi inilah arsitektur jaringan, pola infiltrasi, pleomorfisme nukleus, hingga aktivitas mitosis dinilai oleh patolog untuk menentukan ada atau tidaknya keganasan (WHO, 2019). Pemeriksaan ini bersifat interpretative yang berarti meskipun terdapat kriteria morfologis yang jelas, proses pengambilan keputusan tetap melibatkan penilaian manusia yang dipengaruhi pengalaman dan konteks klinis (Rakha EA et al., 2019).

Dalam praktik sehari-hari, klasifikasi biner antara lesi jinak dan ganas pada kasus yang jelas sering kali memiliki tingkat kesepakatan tinggi antar patolog (Elmore et al., 2015). Namun demikian, variabilitas antar-pemeriksa tetap dilaporkan dalam konteks lesi borderline atau proliferatif atipikal (Elmore et al., 2015). Hal ini menunjukkan bahwa meskipun histopatologi merupakan gold standard (WHO, 2019), reliabilitas tetap menjadi isu metodologis yang relevan dalam penelitian berbasis citra.

Diagnosis histopatologi merupakan hasil dari rangkaian proses panjang mulai dari biopsi atau potongan jaringan hasil reseksi, selanjutnya dilakukan fiksasi dalam cairan fiksatif, prosesing jaringan hingga pemotongan, hingga pewarnaan serta pembacaan dan interpretasi oleh patolog. Proses Panjang ini tentunya juga tidak luput dari terjadinya resiko kerusakan jaringan atau autolisis, sehingga muncul artefak. Adanya aberasi kromatik dan distorsi spasial pun jga mengakibatkan munculnya artefak teknis

mikroskopi, Sedangkan pada tahap pencitraan atau pembuatan foto/citra digital pun dapat muncul artefak digital akibat kesalahan segmentasi, karena preparate histopatologi Adalah representasi dua dimensi dari jaringan yang berstruktur tiga dimensi, maka informasi yang diperoleh pun hanya sebagian dari kompleksitas biologis yang sesungguhnya (Galton et al., 2017).



Gambar 2.1 Contoh kesalahan dari berbagai tahapan alur kerja *imaging*. Kesalahan Level 0: (a) Tampak "pulau" pada epitel yang muncul akibat arah pemotongan yang tidak ortogonal. (b) Jaringan yang terlipat. Kesalahan Level 1: (c) Kiri: latar belakang tidak putih dan pencahayaan tidak merata; tengah: hanya latar belakang; kanan: gambar dikoreksi menggunakan (sampel/bidang terang)×255 untuk setiap warna. (d) Kiri: aberasi kromatik; kanan: gambar dikoreksi menggunakan transformasi kesamaan. (e) Gambar yang dihasilkan oleh paparan 1 menit pada bidang yang tidak diterangi, menunjukkan piksel "panas". Kesalahan Level 2: (f) Kiri, gambar awal yang diambil; tengah: gambar yang kurang tersegmentasi, menggunakan metode histogram "kesalahan minimum"; kanan: segmentasi yang lebih baik menggunakan metode gradien regional (Landini et al., 2017).

Digital Pathology dan Keterbatasannya

Patologi digital (DP) semakin populer di bidang kedokteran, namun menghadapi berbagai keterbatasan yang cukup signifikan. Sebuah studi oleh Snead dkk. (2016) menunjukkan bahwa meskipun ada kemajuan, kualitas gambar digital pada perbesaran tinggi masih kalah bila dibandingkan dengan yang dicapai oleh mikroskop cahaya tradisional, terutama dalam mendeteksi kondisi tertentu seperti Helicobacter pylori, karsinoma sel kecil, dan peradangan granulositik. Dalam beberapa kasus, mencapai akurasi optimal membutuhkan pemindaian resolusi sangat tinggi pada perbesaran ×60 lensa obyektif (Snead et al., 2016). Selain itu, tantangan yang terkait dengan implementasi patologi digital digarisbawahi oleh penelitian dari Jahn dkk. (2020) dan Lujan dkk. (2022), yang mencatat masalah seperti kebutuhan penyimpanan yang besar, yang dapat mencapai tingkat terabyte, serta kebutuhan untuk memindai ulang sekitar 3% kasus karena kesulitan teknis seperti silau, masalah fokus, dan kesalahan pemindaian. Masalah teknis ini, dikombinasikan dengan biaya tinggi dan hambatan manajerial, menghadirkan rintangan signifikan bagi adopsi patologi digital secara luas (Jahn et al., 2020; Lujan et al., 2022).

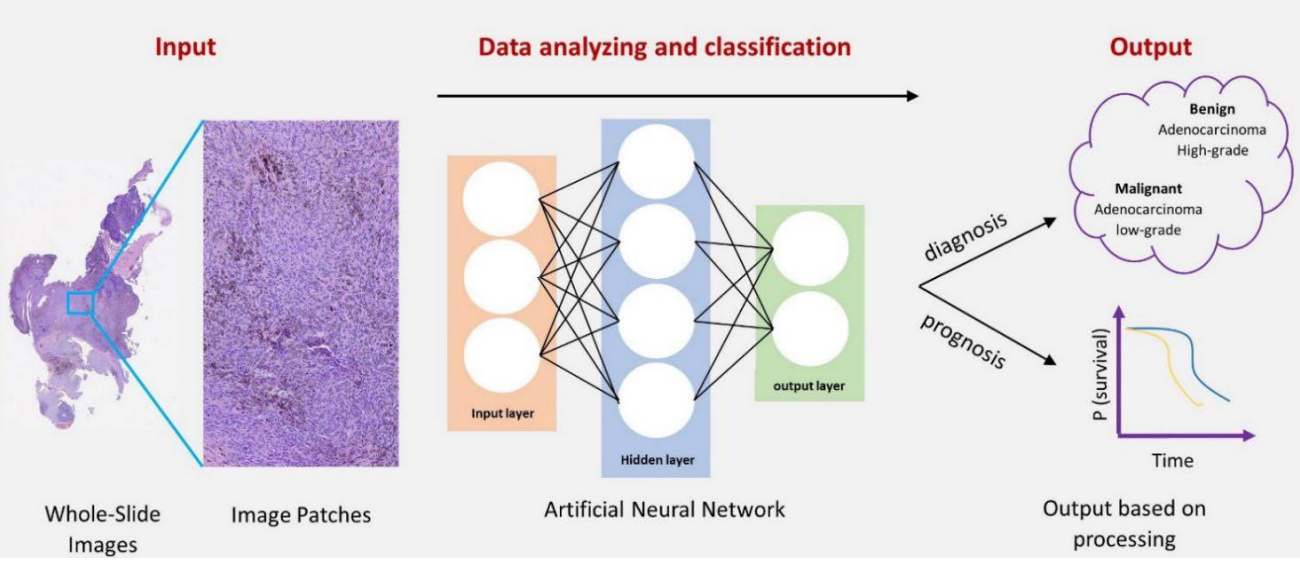
Dalam ranah AI dan segmentasi digital, prosesnya seringkali penuh dengan tantangan, terutama selama tahap kritis ekstraksi fitur. Kesalahan pada tahap ini dapat mengakibatkan segmentasi berlebihan atau kurang, yang pada gilirannya dapat menyebabkan kesalahan klasifikasi data. Masalah ini cukup mengkhawatirkan karena sistem yang dimaksud dapat menghasilkan model histologis yang, meskipun "secara teoritis masuk akal" namun gagal mencerminkan kompleksitas realitas biologis secara akurat. Galton dkk. (2017) juga menekankan pentingnya untuk tidak bergantung pada perangkat

lunak tersebut dengan "kepercayaan buta" dan menyoroti perlunya pengawasan dan evaluasi yang berkelanjutan oleh para ahli patologi untuk memastikan bahwa interpretasi dan keputusan yang dilakukan dilapangan cukup akurat (Galton et al., 2017).

Pengembangan Alat bantu Diagnosis Histopatologis berbasis AI

Selama 10 hingga 15 tahun terakhir, pengembangan alat diagnostik histopatologi berbasis AI, khususnya yang memanfaatkan pembelajaran mendalam (deep learning), telah berkembang pesat, bertepatan dengan pengembangan digitalisasi gambar slide utuh atau *whole slide imaging* (WSI). Meskipun inovasi ini tidak dirancang untuk menggantikan ahli patologi, tujuannya adalah untuk meningkatkan akurasi dan konsistensi diagnostik, mengurangi beban kerja, dan memberikan kuantifikasi temuan secara objektif. Selain itu, inovasi ini berupaya mengintegrasikan data morfologi dan molekuler, menawarkan pendekatan yang lebih komprehensif untuk membantu proses diagnosis. Terlepas dari kinerja AI yang menjanjikan dan tinggi dalam berbagai tugas spesifik, namun perjalanan menuju adopsi klinis masih menghadapi tantangan signifikan, termasuk masalah validasi, regulasi, dan kebutuhan generalisasi di berbagai pengaturan klinis.

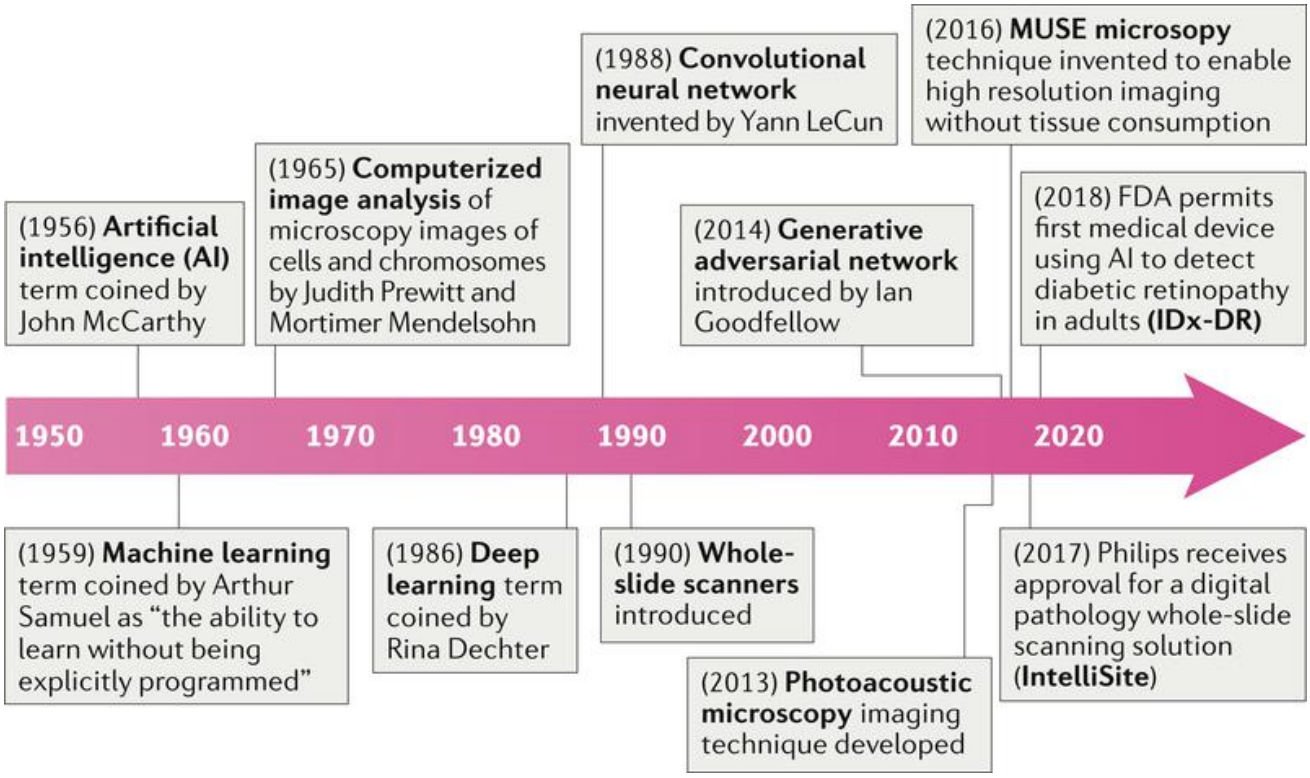
Kemajuan terbaru dalam kecerdasan buatan (AI) telah secara signifikan mengubah patologi tumor, seperti yang disoroti oleh Ahmed et al. (2022). Aplikasi AI kini mencakup hampir setiap aspek diagnostik tumor, termasuk deteksi kondisi seperti metastasis kanker payudara, klasifikasi tumor sebagai jinak atau ganas, dan penentuan tingkat dan stadium kanker. Lebih lanjut, AI telah terbukti mampu mengidentifikasi biomarker dan atribut patologis, memprediksi prognosis dan hasil pasien, dan bahkan menyimpulkan mutasi genetik dari gambar histologi. Yang perlu diperhatikan, studi yang menggunakan Jaringan Saraf Konvolusional (CNN) telah melaporkan akurasi yang luar biasa, dengan beberapa model diagnostik mencapai sensitivitas keseluruhan melebihi 95% (Ahmed et al., 2022).



Gambar 2.2 Gambaran umum proses pembelajaran mendalam (*deep learning*) dalam patologi. Pertama, *Whole Slide Images* (WSI) diambil dari slide spesimen asli. Selanjutnya proses analisis dengan *Artificial Neural Network* (ANN), dan terakhir diperoleh luaran berupa diagnosis atau prognosis yang didasarkan dari klasifikasi dan fitur yang dipilih.

Potensi AI dalam membantu proses diagnosis jauh melampaui pembacaan morfologi histopatologi tradisional. Mereka menggarisbawahi pentingnya mengintegrasikan berbagai teknologi diagnostik yang kompleks seperti genomik, proteomik, imunostaining multipleks, dan data klinis, untuk memperoleh pemahaman yang lebih mendalam tentang profil tumor. Pendekatan komprehensif ini

bertujuan untuk memfasilitasi terapi personalisasi, menggeser AI dari peran sebelumnya sebagai "pengklasifikasi gambar" yang sederhana menjadi alat prediksi multimodal yang canggih. Kemajuan tersebut menunjukkan terjadinya pergeseran transformatif dalam pemanfaatan AI kearah terapi personalisasi pada kasus onkologi atau precision oncology (Bera et al., 2019).



Gambar 2.3. Perjalanan Sejarah Patologi Komputasi

Kemajuan teknologi telah membuat digitalisasi gambar slide utuh (WSI) membantu efisiensi dan menyederhanakan alur kerja di laboratorium patologi di seluruh dunia, dalam dua dekade terakhir. Kondisi ini mendorong dibentuknya perpustakaan slide digital yang cukup ekstensif, terutama *The Cancer Genome Atlas* (Tomczak et al., 2015), yang memberi akses kepada para peneliti terhadap kumpulan data gambar patologi yang telah dikurasi dan dikaitkan dengan informasi klinis dan genomik, sehingga mendorong peningkatan yang sangat berarti pada studi tentang kecerdasan artifisial (AI) dalam patologi dan onkologi digital (Gutman et al., 2013).

Namun demikian, penerapan teknologi ini masih riskan terutama karena masalah generalisasi sehingga menjadi salah satu isu paling mendesak dalam ranah model pembelajaran mendalam (DL), khususnya pada lingkup kesehatan. Pocevičiūtė (2023) menyoroti beberapa tantangan utama yang dihadapi berbagai model ini, seperti kinerja yang seringkali tidak konsisten di berbagai rumah sakit, adanya penurunan akurasi pada kelompok pasien tertentu, serta kinerja model yang fluktuatif dari waktu ke waktu. Bahlan banyak model tidak memiliki kemampuan untuk memberikan skor kepercayaan, sehingga para klinisi tidak mengetahui keandalan prediksi yang dihasilkan. Untuk mengatasi tantangan ini, para peneliti secara aktif mengembangkan solusi seperti deteksi sampel di luar distribusi atau *out-of-distribution* (OOD), teknik estimasi ketidakpastian, dan metode validasi pada berbagai layanan kesehatan. Selain itu, *unsupervised frameworks* masih terus dieksplorasi untuk memprediksi kapan terjadinya penurunan kinerja sebelum model-model ini diimplementasikan dalam praktik klinis. Pergeseran fokus ini menunjukkan bahwa pertanyaan kritisnya bukan lagi apakah AI dapat mencapai akurasi, tetapi apakah AI dapat mempertahankan akurasi tersebut dalam kompleksitas dunia nyata (Pocevičiūtė, 2023).

Cheng dkk. (2021) mengidentifikasi bahwa kebutuhan akan infrastruktur digital yang kuat sangat penting untuk memperkuat implementasi digitalisasi patologi, termasuk kemampuan patologi digital penuh, solusi penyimpanan data yang memadai, serta integrasi sistem laboratorium IT yang memadai. Selain itu, tata ruang regulasi juga menghadirkan tantangan yang dapat menghambat laju inovasi. Lebih lanjut, mendapatkan *ground truth* yang akurat dari patolog untuk membangun algoritma AI sangat penting, sehingga pendekatan baru seperti anotasi *crowdsourcing*, memanfaatkan luaran klinis sebagai tolok ukur, dan menggunakan teknik pembelajaran *semi-supervised* dan *unsupervised*. Terakhir, menumbuhkan kepercayaan dan penerimaan dalam tataran klinis sangat penting, termasuk pentingnya *explainable* AI yang dapat membantu memberikan pejelasan bagaimana AI mendapatkan kesimpulan klinis ataupun memberikan rekomendasi tertentu bagi klinisi, keterlibatan aktif ahli patologi dalam proses pengembangan, serta keterbukaan dalam pembiayaan (Cheng et al., 2021).

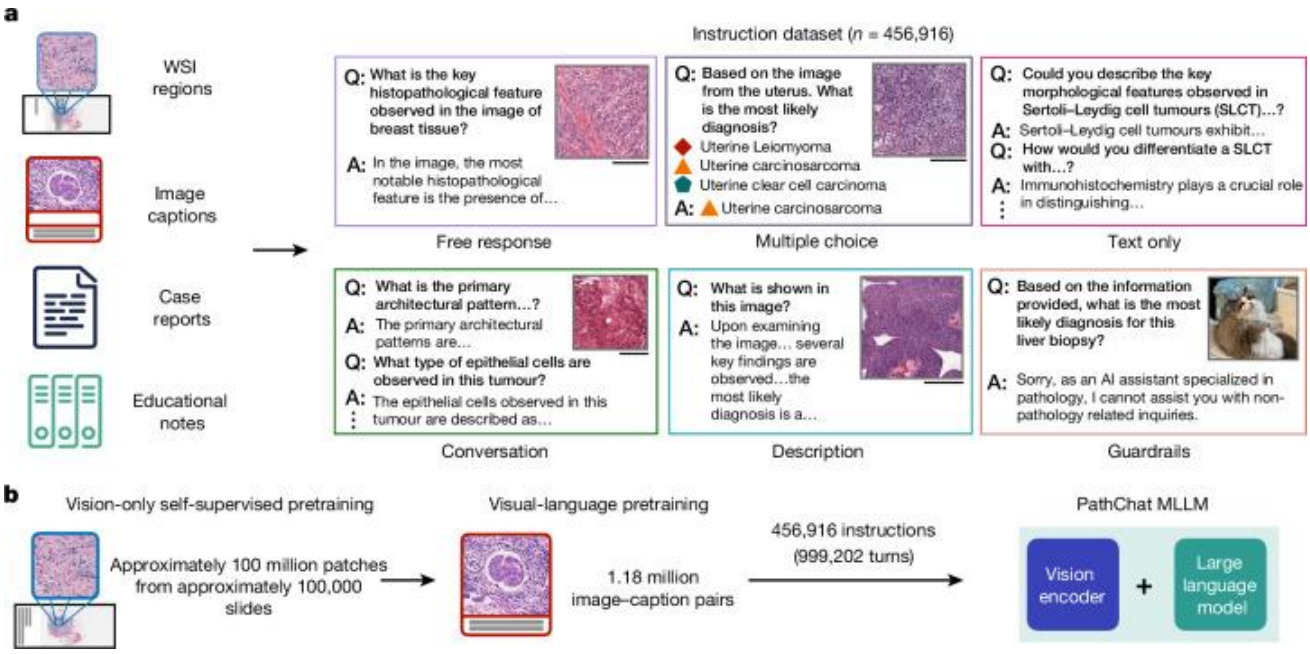
Vision-Language Models (VLMs) dalam Patologi

Di bidang patologi, *Vision-Language Models* (VLM) telah muncul sebagai alat multimodal yang mengintegrasikan gambar histologi, seperti WSI, dengan berbagai bentuk teks patologi, termasuk laporan ekspertise, keterangan morfologi, dan instruksi klinis. Tidak seperti model berbasis visual tradisional yang hanya mengandalkan analisis gambar, VLM ini dirancang untuk mempelajari representasi morfologi visual bersamaan dengan terminologi diagnostik, membangun hubungan yang bermakna antara deskripsi tekstual dan fitur histologis. Pendekatan yang inovatif ini menyerupai cara kerja ahli patologi dalam melakukan tugas di setting klinik, mulai dari mengamati slide dan secara bersamaan menghubungkannya dengan konsep bahasa medis yang relevan. Periode tahun 2023 hingga 2025 dimulainya pergeseran signifikan di bidang ini, yang ditandai dengan transisi dari model spesifik tugas ke model multimodal fundamental, membuka jalan bagi kemajuan dalam akurasi dan efisiensi diagnostik.

Model histopatologi tradisional terutama bergantung pada data gambar, yang seringkali menimbulkan tantangan seperti keterbatasan label medis, yang biasa disebut sebagai kelangkaan label. Namun, laporan patologi kaya akan informasi diagnostik, dan proses penalaran ahli patologi pada dasarnya bersifat multimodal. Menyadari keterbatasan ini, Lu dkk. (2023) memperkenalkan CONCH (Contrastive Learning from Captions for Histopathology), sebuah model bahasa visual inovatif yang dibangun dari lebih dari 1,17 juta pasangan gambar-keterangan. Model inovatif ini tidak terbatas pada satu tugas; melainkan, ia bersifat agnostik terhadap tugas, menunjukkan fleksibilitas untuk diterapkan di berbagai aplikasi, termasuk klasifikasi, segmentasi, pengambilan, dan pemberian keterangan, yang seringkali hanya membutuhkan sedikit atau tanpa penyesuaian. Pengembangan CONCH menyoroti potensi besar penggunaan keterangan morfologis sebagai sinyal pemantauan yang kuat untuk pelatihan awal, yang pada akhirnya menjembatani kesenjangan antara data gambar dan narasi diagnostik yang kaya dalam patologi (Lu, Chen, Williamson, Chen, Liang, et al., 2024).

Salah satu contoh model AI multimodal yang menonjol adalah TITAN, yang diperkenalkan oleh Ding dkk. pada tahun 2024. Model dasar multimodal ini telah dilatih sebelumnya pada kumpulan data besar yang terdiri dari 335.645 gambar WSI utuh dan menggunakan pembelajaran mandiri berbasis visi untuk meningkatkan performanya. Pelatihan TITAN juga menyelaraskan dengan laporan ekspertise patolog yaitu berupa 423.122 keterangan sintetis, sehingga menghasilkan kinerja yang baik tanpa perlu *fine-tuning* dalam tugas-tugas tertentu. Performanya cukup superior pada beberapa bidang seperti klasifikasi few-shot dan zero-shot, pengenalan kanker yang langka, pengambilan data lintas modal, dan bahkan pembuatan laporan patologi (Ding et al., 2025). Kunci keberhasilan TITAN terletak pada pendekatan pra-pelatihan dua tahap yaitu transisi dari keterangan morfologi histopatologi menjadi laporan klinis, dan kemampuannya dalam generalisasi di berbagai skenario klinis, yang menjadikannya berbeda dan superior dibandingkan model-model sebelumnya di bidang ini.

PathChat adalah asisten AI multimodal inovatif yang dirancang khusus untuk patologi. Ia telah dilatih sebelumnya menggunakan kumpulan data yang mengesankan berisi 100 juta gambar histologi dan mencakup 1,18 juta pasangan gambar-keterangan, yang selanjutnya disempurnakan dengan lebih dari 250.000 instruksi bahasa visual. Evaluasi kinerjanya menunjukkan akurasi diagnostik sebesar 87% pada pertanyaan pilihan ganda berbasis kasus publik, dan para ahli patologi menilai respons terbuka yang diberikannya lebih baik dibandingkan dengan model tujuan umum. Seperti yang disoroti dalam sebuah studi oleh Lu dkk. pada tahun 2023, PathChat mencontohkan evolusi dari model klasifikasi tradisional ke asisten AI interaktif yang melibatkan keahlian manusia, menandai kemajuan signifikan di bidang ini (Lu, Chen, Williamson, Chen, Zhao, et al., 2024).



Gambar 2.4 a. Dataset sebanyak 456.916 instruksi dan respons yang sesuai yang mencakup berbagai format (misalnya, percakapan *multi-turn*, pertanyaan pilihan ganda, dan jawaban singkat) dari berbagai sumber. b. Pengembangan asisten VLM AI berbasis MLLM yang dapat berpikir terhadap input visual dan bahasa. (Lu, Chen, Williamson, Chen, Zhao, et al., 2024)

Model vision-language pada citra histopatologi merepresentasikan generasi baru AI patologi yang menggabungkan morfologi visual dan bahasa medis. Model seperti CONCH, TITAN, dan PathChat menunjukkan kemampuan transfer learning, zero-shot classification, dan generasi laporan. Namun, tantangan generalisasi, validasi klinis, dan regulasi tetap menjadi hambatan utama sebelum implementasi rutin di praktik klinis. VLM komersial pun telah hadir untuk membantu patolog, namun masih tahap awal adopsi klinis. Model general seperti GPT-4V telah menunjukkan kemampuan zero-shot yang kompetitif, sementara model domain-spesifik seperti PathChat, CONCH, dan TITAN menunjukkan arah masa depan foundation model patologi multimodal. Saat ini, penggunaan paling realistis adalah sebagai alat bantu diagnostik maupun pendidikan, dan bukan pengganti patolog.

Dalam sebuah studi terbaru yang diterbitkan di Nature Communications, Ferber dkk. (2024) menyelidiki kinerja GPT-4V pada tugas-tugas histopatologi melalui pendekatan pembelajaran kontekstual yang inovatif. Hebatnya, model tersebut tidak memerlukan penyesuaian halus dan hanya bergantung pada contoh yang diberikan dalam perintah. Temuan tersebut mengungkapkan bahwa GPT-4V dapat menyamai atau bahkan mengungguli model khusus pada beberapa tugas penting, termasuk subtipe kanker kolorektal, subtipe polip usus besar, dan deteksi tumor payudara di kelenjar getah bening. Hasil ini menyoroti potensi VLM yang digeneralisasi untuk berkinerja efektif pada

tugas-tugas histopatologi secara langsung, terutama dalam konteks di mana data terbatas (Ferber et al., 2024).

Tabel 2.1 Perbandingan Berbagai AI Asistant Patolog berbasis VLM

Model	**Tipe**	**Status Komersial**	**Kelebihan**	**Keterbatasan**
GPT-4V	Generalist VLM	Sudah komersial	Zero-shot, mudah digunakan	Tidak spesifik patologi
PathChat	Domain-specific VLM	Pra-komersial	Assistant interaktif	Belum luas validasi klinis
CONCH	Foundation VLM	Riset → potensi spin-off	Transferable, multimodal	Butuh integrasi klinis
TITAN	Slide-level foundation VLM	Riset	Strong zero-shot & retrieval	Kompleks dan butuh data besar

Berbagai sistem asisten patologi virtual dan manajemen laboratorium (VLM) komersial saat ini sangat efektif sebagai opini kedua informal dan sumber daya berharga untuk pendidikan dan pelatihan. Sistem ini dapat memfasilitasi penyusunan laporan pendahuluan dan membantu pengumpulan data lintas kasus, menjadikannya alat yang berharga bagi ahli patologi yang mencari efisiensi. Namun, sistem ini belum ideal dan tidak direkomendasikan untuk membuat diagnosis definitif tanpa pengawasan, apalagi dalam kasus yang langka dan kompleks. Lebih lanjut, penggunaan regulasi penuhnya masih memerlukan validasi lokal, yang menggarisbawahi pentingnya keahlian manusia dalam patologi.

BAB 3. METODOLOGI PENELITIAN

3.1. Waktu dan Tempat Penelitian

Penelitian dilakukan pada periode November 2025 sampai Juli 2026. Penelitian ini direncanakan berlangsung selama ±9 bulan, dengan rincian sebagai berikut:

- Bulan 1–2 (November-Desember 2025): Studi literatur dan penyusunan proposal
- Bulan 3 (Januari 2026): Kurasi dataset dan koleksi sampel di RS
- Bulan 4–5 (Februari-Maret 2026): Pembuatan foto mikroskopi, seleksi dan standardisasi citra
- Bulan 6-7 (April-Mei 2026): Proses inferensi menggunakan model AI
- Bulan 8 (Juni 2026): Analisis statistik
- Bulan 9 (Juli 2026): Penyusunan luaran: poster, HKI, manuskrip dan submit publikasi

Tempat penelitian meliputi:

1. Laboratorium Patologi Anatomi RSUD Prof. Dr. Soekandar Mojosari, Mojokerto, Jawa Timur untuk pengumpulan slide histopatologi.
2. Laboratorium Biomedik 2 Fakultas Kedokteran Universitas Trisakti untuk kurasi dan validasi citra internal.
3. Analisis komputasi dilakukan menggunakan komputer pribadi dengan akses API model vision-language komersial. Dataset publik BreakHis diakses melalui repositori resmi.

3.2. Metode Penelitian

3.2.1. Desain penelitian

Penelitian ini adalah studi observasional analitik dengan pendekatan benchmarking performa model vision-language komersial dalam klasifikasi kanker payudara pada citra mikroskopi histopatologi. Pendekatan yang digunakan adalah zero-shot evaluation, di mana model tidak dilatih ulang pada dataset penelitian. Unit analisis utama adalah pasien atau *patient-level analysis*.

3.2.2. Ukuran Sampel

Perhitungan ukuran sampel untuk studi akurasi diagnostik dilakukan berdasarkan rumus yang dikemukakan oleh Buderer (1996) yang mempertimbangkan estimasi sensitivitas dan spesifisitas dengan interval kepercayaan tertentu. Rumus yang digunakan adalah (Fenn Buderer, 1996):

Untuk Sensitifitas: $$N = \frac{Z^2 \times Se(1 - Se)}{d^2 \times Prev}$$

Untuk Spesifisitas: $$N = \frac{Z^2 \times Sp(1 - Sp)}{d^2 \times (1 - Prev)}$$

Karena penelitian ini Adalah benchmarking model AI, maka kita dapat menggunakan asumsi moderat:

- Sensitivitas diharapkan = 0,85
- Spesifisitas diharapkan = 0,85
- Margin error = 0,07 (dipilih karena penelitian ini sebagai studi awal atau benchmarking)
- CI 95% → Z = 1,96
- Prevalensi diperkirakan = 0,5 (untuk mendapatkan dataset seimbang)

Dengan asumsi sensitivitas dan spesifisitas yang diharapkan sebesar 85%, tingkat kepercayaan 95% (Z = 1,96), margin of error 7%, dan prevalensi penyakit 50%, diperoleh estimasi kebutuhan minimal sekitar 200 pasien secara keseluruhan.

Penelitian ini menggunakan pendekatan *consecutive case inclusion within a predefined study period*, yaitu seluruh kasus yang memenuhi kriteria inklusi dalam periode penelitian akan dimasukkan secara berturut-turut tanpa dilakukan pemilihan selektif tambahan.
Pendekatan ini dipilih karena:

1. Dataset publik BreakHis memiliki jumlah pasien tetap (82 pasien), sehingga seluruh pasien yang tersedia akan diikutsertakan.
2. Kohort internal bergantung pada ketersediaan arsip dokumentasi mikroskopi yang memenuhi kriteria inklusi.
3. Studi ini bersifat observasional evaluatif (benchmarking), bukan uji intervensi atau uji superioritas yang memerlukan perhitungan ukuran sampel berbasis hipotesis primer.

Jadi ukuran sampel internal tidak ditentukan melalui rumus perhitungan, melainkan berdasarkan seluruh kasus yang tersedia dan memenuhi kriteria dalam periode penelitian.
Meskipun demikian, untuk menjaga ketelitian estimasi, seluruh parameter performa diagnostik (sensitivitas, spesifisitas, dan akurasi) akan dilaporkan bersama interval kepercayaan 95%. Lebar interval akan menggambarkan ketepatan estimasi (precision), yang merupakan indikator praktis ukuran sampel terlepas dari rumus statistik sebelumnya. Pendekatan ini memungkinkan penilaian presisi estimasi secara transparan tanpa membatasi inklusi kasus yang tersedia.
Strategi ini dipilih untuk:

- Menghindari bias seleksi akibat eksklusi kasus yang memenuhi kriteria.
- Memaksimalkan pemanfaatan data klinis yang tersedia.
- Memastikan bahwa evaluasi performa model dilakukan pada spektrum kasus yang representatif terhadap praktik klinis sehari-hari.

Kriteria inklusi pasien: dataset image diambil dari slide pasien dengan diagnosis PA yang sudah tegak baik tumor ganas (maligna) atauipun tumor jinak (benign) payudara oleh dokter spesialis patologi anatomik
Kriteria eksklusi: data slide histopatologi hilang, rusak atau terlalu banyak artefak

3.2.3. Sumber Data
Penelitian menggunakan dua kohort:
a. Kohort Publik:

- Diambil dari data publik BreakHis, berasal dari 82 pasien (24 *benign*, dan 58 *malignant*)
- Dipilih 1 gambar perbesaran 100× dan 1 gambar 400× per pasien
- Label benign dikategorikan sebagai non-kanker
- Label malignant dikategorikan sebagai kanker

b. Kohort Internal

- RSUD Prof. Dr. Soekandar Mojosari diperoleh sebanyak X pasien (a *benign*, dan b *malignant*) sesuai ketersediaan slide.
- Dibuat dokumentasi mikroskopi dari slide H&E
- Dipilih 1 gambar perbesaran 100× dan 1 gambar 400× per pasien
- Seleksi gambar menggunakan protokol terstandar untuk meminimalkan bias seleksi

3.2.4. Prosedur Pembuatan dan standarisasi foto/citra histopatologi untuk Kohort Internal

1. Sumber Preparat/slide.
   Citra pada kohort internal berasal dari preparat jaringan payudara yang telah diproses secara rutin di Laboratorium Patologi Anatomi menggunakan pewarnaan Hematoxylin-Eosin (H&E).
2. Prosedur Pembuatan foto/citra.

Citra mikroskopi diambil menggunakan mikroskop cahaya dengan kamera digital terintegrasi. Prosedur pengambilan foto dilakukan dengan ketentuan sebagai berikut:

a. Seluruh preparat mikroskopik diperiksa menggunakan mikroskop Olympus CX 20
b. Foto mikroskopik diambil dengan *extended* kamera indomicro 02 yang dihubungkan dengan mikroskop.
c. Pengambilan gambar dilakukan pada area lesi yang telah ditentukan dengan supervisi patolog.
d. Untuk setiap pasien diambil 1 citra pada perbesaran lensa 100×, dan 1 citra pada 400×.
e. Pengaturan pencahayaan dan white balance dibuat konsisten untuk semua slide.
f. Foto disimpan di *google drive* dalam format JPG/PNG supaya kualitas foto tetap terjaga.

3. Spesifikasi dan fitur penting kamera mikroskop untuk algoritme CNN, deep learning adalah sebagai berikut (García-Rojo, 2016; Neary-Zajiczek et al., 2023; Shmatko et al., 2022):
    a. Resolusi ≥ 2 MP (Full HD), atau minimal 2 MP (1920×1080) untuk memastikan detail morfologi jaringan tertangkap dengan baik.
    b. Sensor CMOS BSI atau Sony IMX dengan sensitivitas tinggi dan noise rendah untuk kualitas gambar optimal.
    c. Format gambar: PNG/TIFF (l*ossless*) untuk menjaga integritas data selama analisis AI.
    d. Dukungan kontrol brightness, exposure, white balance untuk menghasilkan reproduksi warna yang akurat dan kontras yang baik sangat penting untuk membedakan fitur histologis yang relevan.
4. Protokol Pemilihan Area. Untuk meminimalkan bias seleksi maka:
    a. Area seleksi diidentifikasi terlebih dulu oleh patolog.
    b. Dari area tersebut, pemilihan bidang pandang dilakukan secara sistematis (bukan memilih area yang dianggap "paling representatif"), dengan protokol sebagai berikut:
        - Patolog mengidentifikasi area lesi secara umum (malignant/benign).
        - Pada magnifikasi 100×:
            - Slide diposisikan di tepi kiri atas area lesi.
            - Mikroskop digeser dari kiri ke kanan secara horizontal.
            - Bidang pandang pertama yang memenuhi kriteria kualitas teknis (fokus baik, tidak ada artefak berat) dipilih.
        - Pada perbesaran 400×:
            - Kembali ke lokasi yang sama.
            - Ulangi prosedur scanning sistematis.
            - Pilih bidang pandang pertama yang memenuhi kriteria teknis.
    c. Untuk slide preparat dari *core biopsy*, maka bidang pandang dipilih dari area tengah lesi pada core, menghindari tepi jaringan dan artefak.

3.2.5. Prosedur Seleksi Gambar pada Kohort Internal

Untuk setiap pasien:

1. Ditentukan diagnosis akhir berdasarkan laporan ahli patologi (gold standard yaitu setelah pembacaan oleh kedua patolog/setelah konsensus).
2. Dipilih 1 gambar pada magnifikasi 100× dan 1 gambar pada magnifikasi 400×.

3. Seleksi dilakukan menggunakan protokol sistematis (bukan berdasarkan subjektivitas "paling representatif").
4. Resolusi dan format gambar distandarisasi sebelum inferensi.

3.2.6. Model yang Dievaluasi

Model vision-language komersial yang dievaluasi:

- GPT-4 Vision
- Gemini Pro Vision
- Claude Sonnet 4.5

Semua model:

- Digunakan melalui API resmi
- Parameter temperature = 0
- Prompt identik/sama
- Tidak dilakukan *fine-tuning*

3.2.7. Penetapan Gold Standard

Untuk kohort internal, penetapan diagnosis referensi dilakukan melalui pembacaan independen oleh dua dokter spesialis Patologi Anatomi yang berpengalaman.

Prosedur dilakukan sebagai berikut:

1. Setiap kasus dinilai secara independen oleh dua patolog (Patolog A dan Patolog B).
2. Penilaian dilakukan secara *blinded*, tanpa mengetahui hasil pembacaan patolog lainnya maupun hasil klasifikasi model AI.
3. Klasifikasi ditetapkan dalam dua kategori:
    - Kanker payudara (malignant)
    - Non-kanker (benign)

Hasil pembacaan awal dari kedua patolog dibandingkan untuk mengidentifikasi tingkat kesepakatan.

Pada kasus dengan perbedaan pendapat (*discordant cases*), dilakukan diskusi bersama untuk mencapai diagnosis konsensus. Diagnosis hasil konsensus ditetapkan sebagai gold standard final yang digunakan dalam analisis performa model AI. Seluruh kasus tetap dimasukkan dalam analisis, termasuk kasus yang awalnya tidak disepakati, guna menghindari bias seleksi dan mempertahankan representativitas klinis dengan menghindari pengeluaran kasus sulit yang dapat menyebabkan inflasi estimasi performa model.

Untuk kohort publik BreakHis, label diagnosis yang tersedia dalam dataset digunakan sebagai gold standard, tanpa dilakukan pembacaan ulang, karena dataset tersebut telah memiliki label referensi yang ditetapkan oleh tim patolog pengembang dataset.

3.2.8. Analisis *Interobserver Agreement*

Tingkat kesepakatan antara dua patolog pada kohort internal dianalisis menggunakan koefisien Cohen's Kappa (κ).

Nilai kappa dihitung berdasarkan klasifikasi biner (kanker vs non-kanker). Interval kepercayaan 95% untuk nilai kappa juga dihitung untuk menilai presisi estimasi.

Analisis ini bertujuan untuk:

1. Menilai reliabilitas penetapan diagnosis referensi.
2. Memastikan bahwa gold standard yang digunakan memiliki konsistensi antar penilai yang memadai.

Selain kappa antar patolog, dihitung pula nilai Cohen's Kappa antara masing-masing model AI dan gold standard final (konsensus) untuk mengevaluasi tingkat kesesuaian model terhadap diagnosis referensi.

3.3. Metode Analisis

3.3.1. Unit Analisis

Unit analisis utama adalah pasien, dan setiap pasien memiliki dua magnifikasi (100× dan 400×), analisis dilakukan:

1. Level magnifikasi (image-level terkontrol, berpasangan)
2. Level pasien (*aggregated decision*)
3. Level kohort (BreakHis dan Internal)
4. Kohort kombinasi

Analisis mempertimbangkan bahwa dua gambar dari pasien yang sama bersifat berpasangan (paired data).

3.3.2. Penyusunan Confusion Matrix

Untuk masing-masing model dan masing-masing magnifikasi, dibuat confusion matrix 2×2:

Tabel 1. Dummy Table untuk Confusion Matrix

	Gold Standard Kanker	Gold Standard Non-Kanker
Prediksi Kanker	True Positive (TP)	False Positive (FP)
Prediksi Non-Kanker	False Negative (FN)	True Negative (TN)

3.3.3 Perhitungan Parameter Diagnostik

Dihitung parameter berikut:

- Accuracy = (TP + TN) / Total
- Sensitivity (Recall) = TP / (TP + FN)
- Specificity = TN / (TN + FP)
- Precision = TP / (TP + FP)
- F1-score = 2 × (Precision × Recall) / (Precision + Recall)

Interval kepercayaan 95% dihitung untuk sensitivitas dan spesifisitas menggunakan metode Wilson.

3.3.4 Analisis Agreement

Tingkat kesesuaian antara hasil model dan gold standard dihitung menggunakan: Cohen's Kappa
Interpretasi kappa mengikuti kategori standar:

- <0.20: poor
- 0.21–0.40: fair
- 0.41–0.60: moderate
- 0.61–0.80: substantial
- 0.80: almost perfect

3.3.5 Analisis Perbandingan Antar Model

Perbandingan performa antar model dilakukan menggunakan McNemar test yang sesuai untuk data berpasangan dan klasifikasi biner (ganas/jinak atau kanker/non-kanker),

Uji ini digunakan untuk menilai apakah terdapat perbedaan proporsi kesalahan klasifikasi yang signifikan antara dua model pada dataset yang sama.
Perbandingan dilakukan secara berpasangan:

- Model A vs Model B
- Model A vs Model C
- Model B vs Model C

Nilai $p < 0.05$ dianggap signifikan.

3.3.6 Analisis Per Magnifikasi
Karena data berpasangan dalam pasien yang sama, dilakukan:

- McNemar test untuk membandingkan performa model pada 100× vs 400×.
- Evaluasi apakah sensitivitas berbeda signifikan antar magnifikasi.

3.3.7 Analisis Per Kohort
Analisis dilakukan secara terpisah untuk:

1. Kohort BreakHis
2. Kohort internal
3. Combined cohort

Tujuannya untuk mengevaluasi stabilitas performa model antar sumber data.

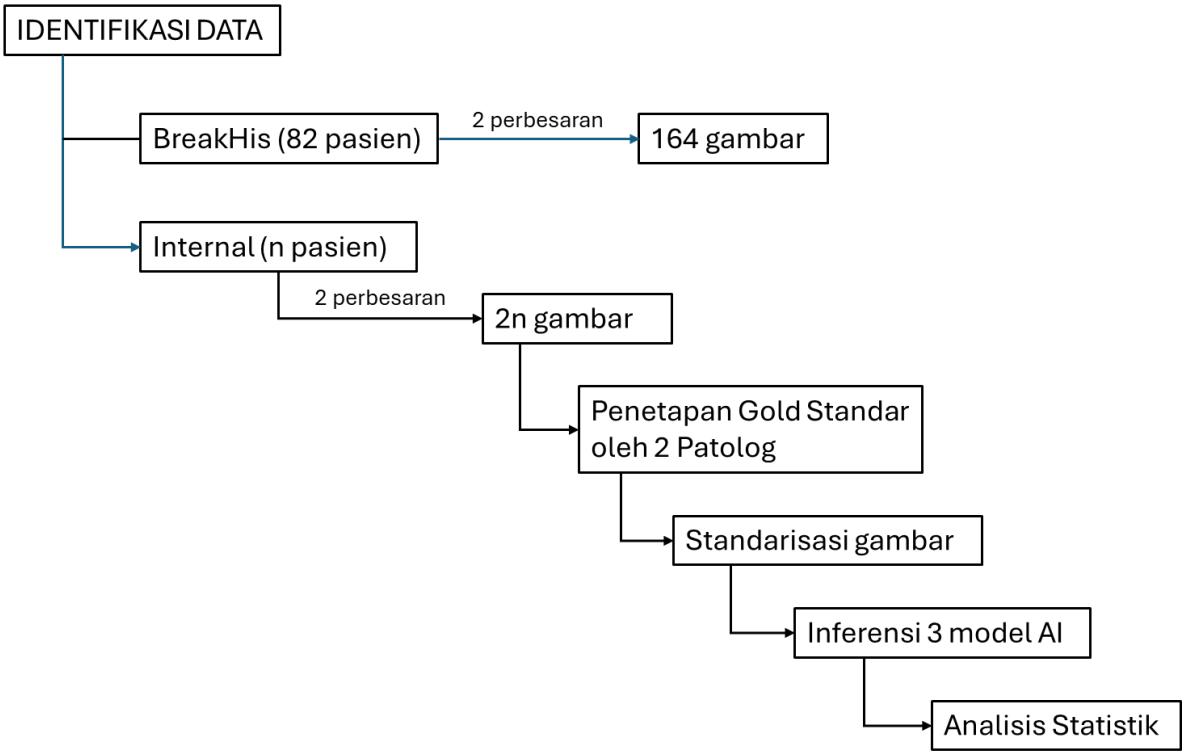
3.3.8 Analisis Tambahan (Eksploratif)
Dilakukan evaluasi terhadap:

- Pola false positive dan false negative
- Karakteristik visual kasus yang sering salah diklasifikasikan

Analisis ini bersifat deskriptif kualitatif.

3.3.9 Perangkat Lunak
Analisis statistik dilakukan menggunakan:

- R atau SPSS
- Tingkat signifikansi ditetapkan pada $\alpha = 0,05$



Gambar 1. Alur Penelitian

Tabel 1. Dummy Table Master Spreadsheet untuk Data Sampel

Patient_ID	Cohort	Specimen_Type	Label	File_100x	File_400x
	BreakHis		BH 1	BH 1_100	BH 1_400
	Internal	Reseksi	IL 1	IL 1_100	IL 1_400
	Internal	Core Biopsy	IL 2	IL 2_100	IL 2_400

3.4. Indikator Capaian Penelitian

ASPEK		**CHECKLIST**
SKALA UNGGULAN	Skala Internasioal	
	Skala Nasional	V
	Skala Lokal	
TOPIK/TEMA RISET	Top Down	V
	Semi Top Down	
	Bottom Up	
SKEMA PENDANAAN	Block Grant	
	Kompetitif	
PELAKSANA RISET	Pusat Penelitian	
	Individu	
	Riset Group	V
SUMBER DANA	Dana Desentralisasi	
	DP2M (30%)	

	Mandiri PT	V
	Kerjasama Luar negeri	
	Sumber Lain-lain	
KEY PERFORMANCE INDICATOR	Jurnal	V
	HKI	V
	Teknologi Tepat Guna	
	S3	
	Seminar	
	Publikasi Internasional	
	Buku Ajar	
	Lain_lain (BUNGA RAMPAI)	
MANAGEMEN PENGELOLAAN	LEMLIT	
	Fakultas	V
	Pusat Penelitian/Studi/Pengkajian	
BUKU PANDUAN	Buku Panduan Penelitian Usakti	V
	Buku Panduan Skim DP2M	
ALOKASI DANA DESENTRALISASI	0-50%	V
	50-75%	
	75-100%	

BAB 4. JADWAL PENELITIAN

No	Kegiatan	2025		2026						
		Nov	Des	Jan	Feb	Maret	April	Mei	Juni	Juli
1	Studi literatur dan penyusunan proposal									
2	Kurasi dataset, koleksi sampel, dan pembacaan oleh patolog.									
3	Pembuatan foto mikroskopi, seleksi dan standardisasi citra.									
4	Proses inferensi dengan model AI.									
5	Analisis statistik									
6	Penyusunan luaran dan laporan lengkap.									

BAB 5. RENCANA ANGGARAN BIAYA (RAB)

No	Komponen	Total
I.	**Biaya Langsung**	
A	Tenaga Ahli dan Tenaga Penunjang	6.400.000
B.1	Biaya bahan habis, peralatan, sewa peralatan	30.540.000
B.2	Perjalanan dan transport lokal	5.310.000
	Sub Total	42.250.000
II.	**Biaya Tidak Langsung**	
A	Laporan, Seminar, dan Publikasi	0
B	Forum Group Disscusion	0
	Sub Total	0
	TOTAL	**42.250.000**

DAFTAR PUSTAKA

Ahmed, A. A., Abouzid, M., & Kaczmarek, E. (2022). Deep Learning Approaches in Histopathology. In *Cancers* (Vol. 14, Number 21). MDPI. https://doi.org/10.3390/cancers14215264

Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Nikšić, M., Bonaventure, A., Valkov, M., Johnson, C. J., Estève, J., Ogunbiyi, O. J., Azevedo e Silva, G., Chen, W. Q., Eser, S., Engholm, G., Stiller, C. A., Monnereau, A., Woods, R. R., Visser, O., … Lewis, C. (2018). Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *The Lancet*, *391*(10125). https://doi.org/10.1016/S0140-6736(17)33326-3

Allison, K. H., Hammond, M. E. H., Dowsett, M., McKernin, S. E., Carey, L. A., Fitzgibbons, P. L., Hayes, D. F., Lakhani, S. R., Chavez-MacGregor, M., Perlmutter, J., Perou, C. M., Regan, M. M., Rimm, D. L., Symmans, W. F., Torlakovic, E. E., Varella, L., Viale, G., Weisberg, T. F., McShane, L. M., & Wolff, A. C. (2020). Estrogen and progesterone receptor testing in breast cancer: ASCO/CAP guideline update. *Journal of Clinical Oncology*, *38*(12). https://doi.org/10.1200/JCO.19.02309

Arnold, M., Morgan, E., Rumgay, H., Mafra, A., Singh, D., Laversanne, M., Vignat, J., Gralow, J. R., Cardoso, F., Siesling, S., & Soerjomataram, I. (2022). Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast*, *66*. https://doi.org/10.1016/j.breast.2022.08.010

Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nature Reviews Clinical Oncology*, *16*(11). https://doi.org/10.1038/s41571-019-0252-y

Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Werneck Krauss Silva, V., Busam, K. J., Brogi, E., Reuter, V. E., Klimstra, D. S., & Fuchs, T. J. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, *25*(8). https://doi.org/10.1038/s41591-019-0508-1

Cheng, J. Y., Abel, J. T., Balis, U. G. J., McClintock, D. S., & Pantanowitz, L. (2021). Challenges in the Development, Deployment, and Regulation of Artificial Intelligence in Anatomic Pathology. In *American Journal of Pathology* (Vol. 191, Number 10). https://doi.org/10.1016/j.ajpath.2020.10.018

Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature Medicine*, *24*(10). https://doi.org/10.1038/s41591-018-0177-5

Ding, T., Wagner, S. J., Song, A. H., Chen, R. J., Lu, M. Y., Zhang, A., Vaidya, A. J., Jaume, G., Shaban, M., Kim, A., Williamson, D. F. K., Robertson, H., Chen, B., Almagro-Pérez, C., Doucet, P., Sahai, S., Chen, C., Chen, C. S., Komura, D., … Mahmood, F. (2025). A multimodal whole-slide foundation model for pathology. *Nature Medicine*, *31*(11). https://doi.org/10.1038/s41591-025-03982-3

Elmore, J. G., Longton, G. M., Carney, P. A., Geller, B. M., Onega, T., Tosteson, A. N. A., Nelson, H. D., Pepe, M. S., Allison, K. H., Schnitt, S. J., O'Malley, F. P., & Weaver, D. L. (2015). Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA - Journal of the American Medical Association*, *313*(11). https://doi.org/10.1001/jama.2015.1405

Fenn Buderer, N. M. (1996). Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Academic Emergency Medicine*, *3*(9). https://doi.org/10.1111/j.1553-2712.1996.tb03538.x

Ferber, D., Wölflein, G., Wiest, I. C., Ligero, M., Sainath, S., Ghaffari Laleh, N., El Nahhas, O. S. M., Müller-Franzes, G., Jäger, D., Truhn, D., & Kather, J. N. (2024). In-context learning enables multimodal large language models to classify cancer pathology images. *Nature Communications* , *15*(1). https://doi.org/10.1038/s41467-024-51465-9

Galton, A., Fouad, S., Landini, G., & Randell, D. (2017). Errors and artefacts in histopathological imaging. *CEUR Workshop Proceedings*, *2137*.

García-Rojo, M. (2016). International Clinical Guidelines for the Adoption of Digital Pathology: A Review of Technical Aspects. In *Pathobiology* (Vol. 83, Numbers 2–3). https://doi.org/10.1159/000441192

Gautama, W. (2022). Breast Cancer in Indonesia in 2022: 30 Years of Marching in Place. *Indonesian Journal of Cancer*, *16*(1). https://doi.org/10.33371/ijoc.v16i1.920

Gutman, D. A., Cobb, J., Somanna, D., Park, Y., Wang, F., Kurc, T., Saltz, J. H., Brat, D. J., & Cooper, L. A. D. (2013). Cancer digital slide archive: An informatics resource to support integrated in silico analysis of TCGA pathology data. *Journal of the American Medical Informatics Association*, *20*(6). https://doi.org/10.1136/amiajnl-2012-001469

Hartmann, L. C., Sellers, T. A., Frost, M. H., Lingle, W. L., Degnim, A. C., Ghosh, K., Vierkant, R. A., Maloney, S. D., Pankratz, V. S., Hillman, D. W., Suman, V. J., Johnson, J., Blake, C., Tlsty, T., Vachon, C. M., Melton, L. J., & Visscher, D. W. (2005). Benign Breast Disease and the Risk of Breast Cancer. *New England Journal of Medicine*, *353*(3). https://doi.org/10.1056/nejmoa044383

Jahn, S. W., Plass, M., & Moinfar, F. (2020). Digital pathology: Advantages, limitations and emerging perspectives. In *Journal of Clinical Medicine* (Vol. 9, Number 11). https://doi.org/10.3390/jcm9113697

Khomeini. (2024). PROFIL PASIEN TUMOR JINAK (FAM) DI RSUD M. ZEIN PAINAN PADA TAHUN 2023. *Journal of Public Health Science*, *1*(4). https://doi.org/10.70248/jophs.v1i4.1905

Landini, G., Randell, D. A., Fouad, S., & Galton, A. (2017). Automatic thresholding from the gradients of region boundaries. *Journal of Microscopy*, *265*(2). https://doi.org/10.1111/jmi.12474

Lu, M. Y., Chen, B., Williamson, D. F. K., Chen, R. J., Liang, I., Ding, T., Jaume, G., Odintsov, I., Le, L. P., Gerber, G., Parwani, A. V., Zhang, A., & Mahmood, F. (2024). A visual-language foundation model for computational pathology. *Nature Medicine*, *30*(3). https://doi.org/10.1038/s41591-024-02856-4

Lu, M. Y., Chen, B., Williamson, D. F. K., Chen, R. J., Zhao, M., Chow, A. K., Ikemura, K., Kim, A., Pouli, D., Patel, A., Soliman, A., Chen, C., Ding, T., Wang, J. J., Gerber, G., Liang, I., Le, L. P., Parwani, A. V., Weishaupt, L. L., & Mahmood, F. (2024). A multimodal generative AI copilot for human pathology. *Nature*, *634*(8033). https://doi.org/10.1038/s41586-024-07618-3

Lujan, G., Li, Z., & Parwani, A. V. (2022). Challenges in implementing a digital pathology workflow in surgical pathology. In *Human Pathology Reports* (Vol. 29). https://doi.org/10.1016/j.hpr.2022.300673

Neary-Zajiczek, L., Beresna, L., Razavi, B., Pawar, V., Shaw, M., & Stoyanov, D. (2023). Minimum resolution requirements of digital pathology images for accurate classification. *Medical Image Analysis*, *89*. https://doi.org/10.1016/j.media.2023.102891

Nguyen, A. T., Vuong, T. T. Le, & Kwak, J. T. (2024). Towards a Text-Based Quantitative and Explainable Histopathology Image Analysis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* , *15004 LNCS*. https://doi.org/10.1007/978-3-031-72083-3_48

Pasaribu, E., Setiaji, K., Wahyono, A., & Pasaribu, E. (2020). Clinicopathological and Subtype Breast Cancer in Indonesia : Survival Analysis. *European Journal of Surgical Oncology*, *46*(2). https://doi.org/10.1016/j.ejso.2019.11.132

Pavithra D, Harini R, Hemalatha A, AravindRaj P, Megha D, Pavithra P, & Nirmala E. (2023). Pathophysiological and epidemiological aspects of fibroadenoma of the breast: A review. *World Journal of Biology Pharmacy and Health Sciences*, *15*(1). https://doi.org/10.30574/wjbphs.2023.15.1.0297

Pocevičiūtė, M. (2023). *Generalisation and reliability of deep learning for digital pathology in a clinical setting* (Vol. 2337). https://doi.org/10.3384/9789180753005

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. *Proceedings of Machine Learning Research*, *139*.

Rakha EA, Allison, K. H., Ellis IO, Horii R, Masuda S, & Penalut-Llorca F. (2019). WHO classification of tumours of the breast tumours. *World Health Organization Classification of Tumours*, (4 (der 4. ed.)).

Shmatko, A., Ghaffari Laleh, N., Gerstung, M., & Kather, J. N. (2022). Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. In *Nature Cancer* (Vol. 3, Number 9). https://doi.org/10.1038/s43018-022-00436-4

Singh, D., Vignat, J., Lorenzoni, V., Eslahi, M., Ginsburg, O., Lauby-Secretan, B., Arbyn, M., Basu, P., Bray, F., & Vaccarella, S. (2023). Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO Global Cervical Cancer Elimination Initiative. *The Lancet Global Health*, *11*(2). https://doi.org/10.1016/S2214-109X(22)00501-0

Snead, D. R. J., Tsang, Y. W., Meskiri, A., Kimani, P. K., Crossman, R., Rajpoot, N. M., Blessing, E., Chen, K., Gopalakrishnan, K., Matthews, P., Momtahan, N., Read-Jones, S., Sah, S., Simmons, E., Sinha, B., Suortamo, S., Yeo, Y., El Daly, H., & Cree, I. A. (2016). Validation of digital

pathology imaging for primary histopathological diagnosis. *Histopathology*, *68*(7). https://doi.org/10.1111/his.12879

Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016). A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Transactions on Biomedical Engineering*, *63*(7). https://doi.org/10.1109/TBME.2015.2496264

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209–249. https://doi.org/10.3322/caac.21660

Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. In *Wspolczesna Onkologia* (Vol. 1A). https://doi.org/10.5114/wo.2014.47136

Utamie, H. M., Ningrom, I. C., Handoyo Sakti, Y. B., & Maulana, A. Muh. (2023). Comparison of Pericanalicular and Intracanalicular Fibroadenoma Mammae Growth Pattern on Various Age Range. *MAGNA MEDICA Berkala Ilmiah Kedokteran Dan Kesehatan*, *10*(2). https://doi.org/10.26714/magnamed.10.2.2023.191-198

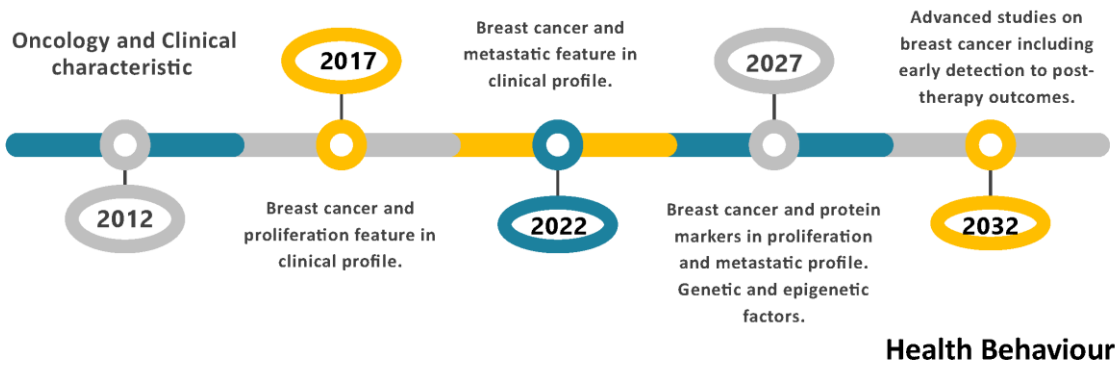
WHO. (2019). Classification of Tumours Editorial Board. *International Agency for Pesearch on Cancer*, *2*(5[a]).

Zeinomar, N., Phillips, K. A., Daly, M. B., Milne, R. L., Dite, G. S., MacInnis, R. J., Liao, Y., Kehm, R. D., Knight, J. A., Southey, M. C., Chung, W. K., Giles, G. G., McLachlan, S. A., Friedlander, M. L., Weideman, P. C., Glendon, G., Nesci, S., Andrulis, I. L., Buys, S. S., … Terry, M. B. (2019). Benign breast disease increases breast cancer risk independent of underlying familial risk profile: Findings from a Prospective Family Study Cohort. *International Journal of Cancer*, *145*(2). https://doi.org/10.1002/ijc.32112

Zhang, Y., Wang, X., Xu, Z., Yu, Q., Yuille, A., & Xu, D. (2020). When radiology report generation meets knowledge graph. *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v34i07.6989

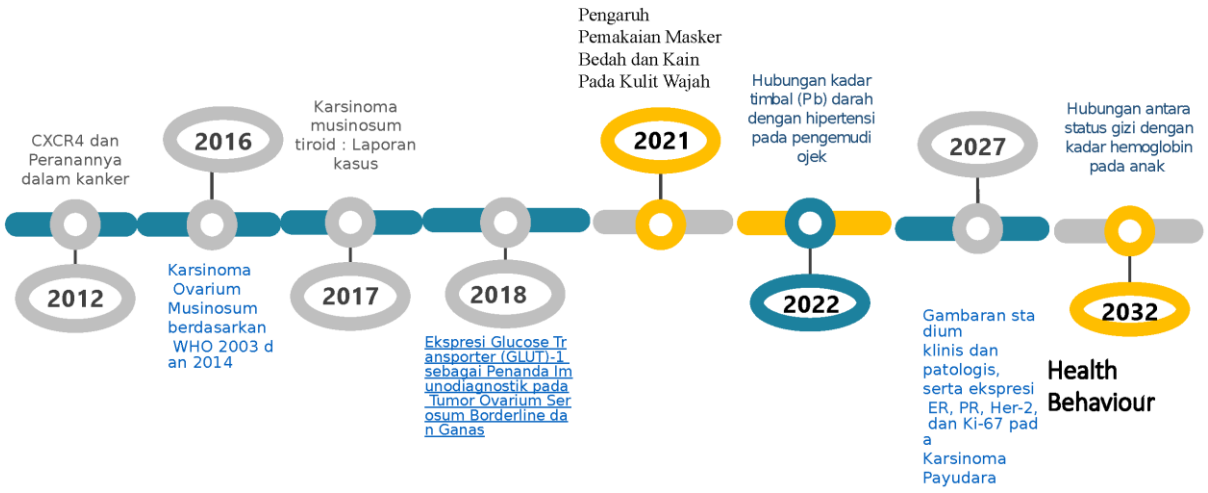
…

LAMPIRAN 1. ROAD MAP PENELITIAN



PETA JALAN PENELITIAN

<dr. Dyah Ayu Woro Setyaningrum, M.Biomed.>, <3507/USAKTI>, <0331057607>



***ROAD MAP* PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT**

<*Florinda Ilona*>, <*3612*>, <*0322107601*>

LAMPIRAN 2. SURAT KESEDIAAN BERKOMITMEN