

Evaluasi Performa Sistem Kuantifikasi Otomatis Imunohistokimia Berbasis Kecerdasan Buatan Untuk Analisis Ekspresi Tumor Marker

Performance Evaluation of an AI-Based Automated Immunohistochemistry Quantification System for Tumor Marker Expression

Cut Indriputri¹, Maulia Hardian Hayati², Rafika³, Yaumul Fachni Tandjungbulu⁴, Reza Maulana⁵

^{1,2,3,4}Jurusan Teknologi Laboratorium Medis, Poltekkes Kemenkes Makassar

⁵Jurusan Sistem Informasi, Fakultas Saintek, UIN Alauddin Makassar

*E-mail korespondensi: indricut@poltekkes-mks.ac.id, 081222449577

ABSTRACT

The evaluation of automated immunohistochemistry (IHC) quantification systems has become increasingly important to ensure accuracy and consistency in tumor biomarker expression analysis. This systematic review aims to assess the performance of AI-based automated IHC quantification systems in both clinical and research settings. A comprehensive literature search was conducted in PubMed, Scopus, and ScienceDirect for studies published between 2015 and 2025. A total of 26 eligible studies were included and qualitatively analyzed in accordance with PRISMA 2020 guidelines. The findings indicate that AI-based IHC quantification systems demonstrate high diagnostic performance, with reported accuracy ranging from 90% to 98%, area under the curve (AUC) values up to 0.96, and correlation coefficients exceeding 0.95 when compared with expert pathologist assessments across key biomarkers such as HER2, Ki-67, ER, PR, P53, and PD-L1. Deep learning models, particularly convolutional neural networks (CNNs) including ResNet and DenseNet architectures, consistently show strong performance in tumor region detection and automated biomarker quantification. Overall, AI improves reproducibility by reducing interobserver variability and enhancing consistency with reference standards, while also standardizing interpretation in borderline cases. Methodological evolution reveals a transition from rule-based digital image analysis to more adaptive deep learning approaches, resulting in improved accuracy and generalizability. However, performance remains influenced by pre-analytical variability, including staining quality, dataset heterogeneity, sample size limitations, and the need for pathologist supervision, particularly in fully automated tumor detection. Despite its strong potential to enhance diagnostic objectivity and efficiency, clinical implementation of AI-based IHC quantification requires further standardization, multicenter validation, and prospective evaluation..

Keywords : *Immunohistochemistry, Artificial Intelligence, Tumor Marker Expression, Automated Quantification, Digital Image Analysis, Deep Learning*

ABSTRAK

Evaluasi kinerja sistem kuantifikasi otomatis imunohistokimia (IHC) semakin penting untuk menjamin akurasi dan konsistensi dalam analisis ekspresi penanda tumor. Tinjauan sistematis ini bertujuan untuk mengevaluasi kinerja sistem kuantifikasi IHC otomatis berbasis kecerdasan buatan dalam konteks klinis dan penelitian. Penelusuran literatur dilakukan pada basis data PubMed, Scopus, dan ScienceDirect untuk publikasi tahun 2015–2025. Sebanyak 26 studi yang memenuhi kriteria inklusi dianalisis secara kualitatif berdasarkan pedoman PRISMA 2020. Hasil menunjukkan bahwa sistem kuantifikasi IHC berbasis kecerdasan buatan memiliki performa diagnostik tinggi dengan akurasi 90–98%, AUC hingga 0,96, serta korelasi dengan ahli patologi yang melebihi 0,95 pada berbagai biomarker utama seperti HER2, Ki-67, ER, PR, P53, dan PD-L1. Model *deep learning* seperti *convolutional neural networks* (CNN), termasuk ResNet dan DenseNet, menunjukkan kemampuan unggul dalam deteksi area tumor dan kuantifikasi ekspresi biomarker secara otomatis. AI juga terbukti meningkatkan reproduksibilitas hasil dengan mengurangi variasi antar-pengamat serta meningkatkan kesesuaian dengan metode referensi, sekaligus menstandarkan interpretasi pada kasus borderline. Perkembangan metode menunjukkan pergeseran dari *digital image analysis* berbasis aturan menuju *deep learning* yang lebih adaptif, sehingga meningkatkan akurasi dan generalisasi. Namun, kinerja sistem masih dipengaruhi oleh variasi pra-analitik, termasuk kualitas pewarnaan, heterogenitas data, ukuran dataset, serta kebutuhan supervisi ahli patologi. Secara keseluruhan, meskipun menunjukkan potensi besar dalam meningkatkan objektivitas dan efisiensi diagnostik, implementasi klinis AI dalam kuantifikasi IHC masih memerlukan standardisasi, validasi multisenter, dan evaluasi prospektif lebih lanjut.

Kata kunci : Imunohistokimia (IHC), Kecerdasan Buatan (AI), Ekspresi Penanda Tumor, Kuantifikasi Otomatis, Analisis Citra Digital, *Deep Learning*

PENDAHULUAN

Imunohistokimia tetap menjadi salah satu pilar utama dalam bidang patologi karena mampu menyediakan data penting untuk mengidentifikasi protein spesifik dalam sampel jaringan. Metode ini berperan krusial dalam menentukan subtipe tumor melalui analisis ekspresi biomarker seperti HER2, ER, PR, dan Ki-67. Selama beberapa dekade, interpretasi hasil IHC sangat bergantung pada penilaian visual oleh ahli patologi terhadap intensitas pewarnaan. Meskipun berbasis keahlian, pendekatan ini tidak terlepas dari potensi bias subjektif serta variasi antar-pengamat (Bankhead *et al.* 2017).

Perkembangan patologi digital telah menghadirkan berbagai perangkat otomatis untuk analisis IHC yang bertujuan meningkatkan objektivitas dalam evaluasi. Platform perangkat lunak seperti QuPath, HALO, ImmunoPath, dan Visiopharm memungkinkan pengukuran yang lebih presisi terhadap rasio sel positif dan kepadatan pewarnaan (Horai *et al.* 2019). Teknologi ini terbukti sangat bermanfaat terutama pada studi berskala besar, di mana perhitungan manual menjadi tidak efisien.

Dalam perkembangan terbaru, integrasi sistem berbasis kecerdasan buatan (*artificial intelligence/AI*), khususnya dengan pendekatan *deep learning*, semakin meningkatkan stabilitas dan konsistensi hasil kuantifikasi dengan kemampuan menganalisis pola pewarnaan yang kompleks (Brattoli *et al.* 2024). Namun demikian, penerapan sistem otomatis secara penuh masih menghadapi sejumlah tantangan, seperti variasi proses pewarnaan antar laboratorium serta belum adanya standar kalibrasi yang seragam, yang dapat memengaruhi validitas analisis (Choi *et al.* 2024).

Meskipun berbagai penelitian telah melaporkan keberhasilan penerapan AI pada analisis IHC, sebagian besar publikasi masih berfokus pada pengembangan algoritma atau evaluasi performa pada biomarker tertentu secara terpisah. Hingga saat ini, masih terbatas tinjauan yang secara komprehensif membandingkan performa berbagai sistem kuantifikasi otomatis berbasis AI, sekaligus mengevaluasi faktor-faktor teknis yang memengaruhi akurasi, reproduksibilitas, dan kesiapan implementasinya dalam praktik patologi diagnostik. Kesenjangan ini menyulitkan penentuan sejauh mana teknologi tersebut benar-benar mampu menggantikan atau melengkapi penilaian konvensional oleh ahli patologi.

Berdasarkan kesenjangan tersebut, tinjauan ini menyajikan evaluasi komprehensif mengenai performa sistem kuantifikasi otomatis imunohistokimia berbasis AI dengan mengintegrasikan bukti terkini mengenai akurasi, konsistensi, keterbatasan teknis, serta tantangan implementasi klinis pada berbagai tumor marker. Berbeda dengan tinjauan sebelumnya yang umumnya membahas aspek algoritma atau aplikasi pada biomarker tertentu, artikel ini menekankan analisis komparatif terhadap kesiapan teknologi untuk diadopsi dalam praktik diagnostik rutin sehingga dapat memberikan dasar ilmiah bagi pengembangan standar penerapan AI dalam patologi digital.

METODE

Desain, tempat dan waktu

Penelitian ini merupakan tinjauan sistematis (*systematic review*) yang bertujuan mengevaluasi performa sistem kuantifikasi imunohistokimia (*immunohistochemistry*, IHC) otomatis berbasis kecerdasan buatan (*artificial intelligence*, AI) dalam analisis ekspresi tumor marker. Proses penyusunan *systematic review* mengacu pada pedoman PRISMA 2020, yang meliputi tahapan identifikasi, penyaringan (*screening*), penilaian kelayakan (*eligibility*), dan inklusi (*included*) terhadap artikel yang memenuhi kriteria.

Penelusuran artikel dilakukan melalui basis data ilmiah internasional, yaitu PubMed, Scopus, dan ScienceDirect. Proses pencarian, seleksi literatur, ekstraksi data, dan analisis dilaksanakan pada periode Januari–April 2026. Artikel yang disertakan merupakan publikasi berbahasa Inggris yang diterbitkan pada rentang tahun 2015–2025.

Jenis dan Cara Pengumpulan Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari artikel penelitian (*original research articles*) yang dipublikasikan dalam jurnal ilmiah internasional. Penelusuran literatur dilakukan pada basis data PubMed, Scopus, dan ScienceDirect terhadap artikel yang diterbitkan pada periode 2015–2025.

Strategi pencarian disusun menggunakan kombinasi kata kunci yang berkaitan dengan kuantifikasi imunohistokimia berbasis kecerdasan buatan, dengan memanfaatkan operator Boolean *AND* dan *OR*. Kata kunci yang digunakan meliputi "*immunohistochemistry*" OR "*IHC*", "*artificial intelligence*" OR "*deep learning*" OR "*machine learning*", "*automated quantification*" OR "*digital image analysis*" OR "*automated scoring*", serta "*tumor marker*" OR "*biomarker expression*". Contoh strategi pencarian yang digunakan adalah:

("immunohistochemistry" OR "IHC") AND ("artificial intelligence" OR "deep learning" OR "machine learning") AND ("automated quantification" OR "digital image analysis" OR "automated scoring") AND ("tumor marker" OR "biomarker expression")

Seleksi artikel dilakukan sesuai pedoman PRISMA 2020 melalui tahapan identifikasi, penyaringan judul dan abstrak, penilaian naskah lengkap, serta penetapan artikel yang memenuhi kriteria inklusi. Kriteria inklusi meliputi penelitian yang menggunakan sampel jaringan manusia, menerapkan sistem kuantifikasi imunohistokimia berbasis kecerdasan buatan, dan melaporkan evaluasi performa sistem, seperti akurasi, sensitivitas, spesifisitas, presisi, F1-score, *area under the curve* (AUC), atau tingkat kesesuaian dengan penilaian ahli patologi. Artikel berupa tinjauan pustaka, editorial, prosiding konferensi, penelitian pada hewan, penelitian yang hanya membahas penilaian manual, maupun penelitian yang tidak menyajikan evaluasi performa sistem dikeluarkan dari analisis.

Berdasarkan proses seleksi tersebut, sebanyak 26 artikel memenuhi seluruh kriteria inklusi dan selanjutnya dianalisis secara kualitatif. Artikel yang diperoleh dari ketiga basis data terlebih dahulu diperiksa untuk menghilangkan duplikasi. Selanjutnya dilakukan penyaringan berdasarkan judul dan abstrak, diikuti dengan penilaian terhadap naskah lengkap sesuai kriteria inklusi dan eksklusi. Artikel yang memenuhi seluruh kriteria kemudian dimasukkan ke dalam proses ekstraksi data dan analisis secara kualitatif.

Pengolahan dan Analisis Data

Data dari artikel yang terpilih dianalisis secara deskriptif dengan meninjau pendekatan metodologis, alur kerja sistem, serta kinerja metode kuantifikasi IHC otomatis. Tahapan analisis mencakup identifikasi proses akuisisi citra, prapemrosesan, segmentasi, serta algoritma penilaian yang digunakan dalam masing-masing studi. Metode deteksi yang dianalisis meliputi teknik *color deconvolution*, segmentasi berbasis ambang (threshold), serta klasifikasi berbasis kecerdasan buatan. Pendekatan *deeplearning* menggunakan *convolutional neural networks* (CNN) untuk ekstraksi fitur dan klasifikasi, sedangkan metode *machine learning* memanfaatkan algoritma seperti *support vector machine* (SVM) dan *random forest*. Validasi kinerja sistem dievaluasi berdasarkan parameter seperti kesesuaian dengan penilaian ahli patologi, variabilitas antar-pengamat, serta analisis *receiver operating characteristic* (ROC). Selain itu, dilakukan perbandingan antara metode manual dan sistem otomatis untuk menilai peningkatan objektivitas, reproduibilitas, dan keandalan diagnostik. Hasil analisis juga mencakup evaluasi kinerja perangkat lunak seperti QuPath, HALO, dan Visiopharm berdasarkan akurasi deteksi, kecepatan analisis, serta kemudahan penggunaan. Secara keseluruhan, analisis difokuskan pada kemampuan sistem dalam mengurangi bias subjektif dan meningkatkan konsistensi hasil antar studi dan antar laboratorium.

HASIL DAN PEMBAHASAN

Penelusuran literatur secara komprehensif dilakukan melalui basis data PubMed, ScienceDirect, dan Google Scholar, yang menghasilkan 52 artikel ilmiah terkait kuantifikasi imunohistokimia (IHC) otomatis. Setelah melalui proses seleksi berdasarkan kriteria inklusi dan eksklusi, sebanyak 26 studi yang dipublikasikan pada periode 2015–2025 dipilih untuk dianalisis lebih lanjut. Studi-studi tersebut berasal dari berbagai negara, dengan kontribusi terbesar dari Amerika Serikat, Inggris, dan Tiongkok.

1. Performa diagnostik sistem AI pada kuantifikasi IHC

Dari 26 penelitian yang dianalisis, sebagian besar melaporkan bahwa sistem kuantifikasi imunohistokimia berbasis AI memiliki performa diagnostik yang tinggi dengan tingkat akurasi berkisar antara 90–98% pada berbagai biomarker dan jenis kanker. Model *deep learning*

seperti *convolutional neural networks* (CNN), ResNet, DenseNet, maupun LadderNet secara konsisten menunjukkan kemampuan yang sangat baik dalam mendeteksi area tumor, mengenali intensitas pewarnaan, dan melakukan kuantifikasi biomarker secara otomatis. Beberapa penelitian bahkan melaporkan nilai AUC hingga 0,96, korelasi dengan ahli patologi melebihi 0,95, serta tingkat kesesuaian terhadap penilaian referensi mencapai 100% pada biomarker tertentu. Temuan tersebut menunjukkan bahwa AI tidak hanya mampu mereplikasi penilaian manual, tetapi juga memberikan evaluasi yang lebih objektif dan konsisten. Keunggulan ini terutama terlihat pada biomarker yang memerlukan penilaian kuantitatif, seperti HER2, Ki-67, ER, PR, IL-24, dan P53, yang selama ini rentan terhadap variasi interpretasi antar-patolog.

2. AI meningkatkan reproduksibilitas dan mengurangi subjektivitas

Salah satu pola yang paling konsisten pada seluruh penelitian adalah kemampuan AI dalam meningkatkan reproduksibilitas hasil pemeriksaan. Hampir seluruh studi menunjukkan bahwa sistem otomatis mampu mengurangi variasi antar-pengamat (*interobserver variability*) yang selama ini menjadi kelemahan utama penilaian imunohistokimia secara manual. Pada evaluasi HER2, penggunaan AI bahkan meningkatkan kesepakatan antar-patolog dari 75% menjadi 83,7%, sedangkan beberapa penelitian lain menunjukkan kesesuaian yang sangat tinggi terhadap konsensus ahli maupun metode referensi seperti FISH dan DISH. Selain itu, beberapa penelitian menunjukkan bahwa analisis digital mampu mengurangi kasus borderline atau equivocal, meningkatkan konsistensi antar laboratorium, serta menghasilkan klasifikasi biomarker yang lebih seragam. Hal ini menunjukkan bahwa AI berpotensi menjadi alat standarisasi dalam patologi digital, khususnya pada pemeriksaan yang memerlukan keputusan terapi berbasis biomarker.

3. Perkembangan algoritma dari *image analysis* menuju *deep learning*

Analisis terhadap seluruh penelitian memperlihatkan adanya perkembangan teknologi yang jelas selama satu dekade terakhir. Penelitian awal (2016–2018) umumnya menggunakan metode *color deconvolution*, *thresholding*, *Gaussian mixture model*, dan teknik segmentasi berbasis morfologi. Metode-metode tersebut sudah mampu meningkatkan objektivitas dibandingkan penilaian manual, tetapi masih sangat bergantung pada parameter yang ditentukan pengguna. Sebaliknya, penelitian yang lebih baru (2023–2025) didominasi oleh pendekatan *deep learning*, yang memungkinkan sistem mempelajari karakteristik morfologi jaringan secara otomatis tanpa memerlukan aturan segmentasi yang kaku. Pendekatan ini menghasilkan akurasi yang lebih tinggi, kemampuan generalisasi yang lebih baik, serta memungkinkan analisis berbagai biomarker dan jenis tumor dalam satu kerangka kerja. Perubahan ini menunjukkan bahwa perkembangan AI pada bidang imunohistokimia telah bergeser dari sekadar digital image analysis menuju sistem cerdas yang semakin mendekati kemampuan interpretasi ahli patologi.

4. Faktor yang memengaruhi performa sistem AI

Walaupun secara umum menunjukkan performa tinggi, hampir semua penelitian melaporkan bahwa akurasi sistem AI masih dipengaruhi oleh berbagai faktor teknis. Variasi protokol pewarnaan, kualitas preparat, jenis antibodi, karakteristik pemindai digital, ukuran dataset pelatihan, serta heterogenitas jaringan merupakan faktor yang paling sering dilaporkan memengaruhi hasil kuantifikasi. Beberapa penelitian juga menemukan bahwa performa sistem menurun ketika deteksi area tumor dilakukan secara otomatis penuh tanpa supervisi ahli patologi, terutama akibat masuknya area non-tumor atau artefak preparat. Temuan ini menunjukkan bahwa keberhasilan implementasi AI tidak hanya ditentukan oleh algoritma yang digunakan, tetapi juga oleh kualitas data masukan dan standarisasi proses pra-analitik maupun analitik.

5. Tantangan implementasi klinis dan arah pengembangan

Meskipun sebagian besar penelitian menunjukkan hasil yang sangat menjanjikan, implementasi AI dalam praktik klinis masih menghadapi sejumlah tantangan. Banyak penelitian menekankan perlunya validasi multisenter, standarisasi protokol pewarnaan, harmonisasi antar perangkat pemindai, serta evaluasi pada populasi pasien yang lebih beragam sebelum sistem dapat digunakan secara rutin. Selain itu, sebagian besar penelitian masih menggunakan desain retrospektif

dengan ukuran dataset yang relatif terbatas sehingga generalisasi hasil masih perlu dibuktikan lebih lanjut.

Secara keseluruhan, sintesis dari 26 penelitian menunjukkan bahwa AI memiliki potensi besar untuk meningkatkan objektivitas, efisiensi, dan reproduksibilitas kuantifikasi imunohistokimia. Namun, adopsi luas dalam praktik diagnostik memerlukan pengembangan standar validasi yang seragam, integrasi dengan alur kerja laboratorium, serta evaluasi prospektif pada kondisi klinis nyata.

Tabel 1. Ringkasan studi yang ditinjau mengenai kinerja metode kuantifikasi imunohistokimia (IHC) otomatis dalam penilaian ekspresi penanda tumor

No	Studi (Penulis, Tahun)	Biomarker	Antibodi / Klon	Sampel / Setting	Metode / Platform	Temuan Utama (Performa)	Catatan / Implikasi
1	(Boyaci <i>et al.</i> 2025)	Ki-67	Klon 30-9	4.313 kanker payudara	Hotspot vs global digital scoring	Global scoring meningkatkan konsistensi antar laboratorium	Kategori intermediate lebih luas memengaruhi keputusan klinis
2	(Hou <i>et al.</i> 2025)	ALK	BP6165 vs D5F3	87 biopsi paru	LYNX480 PLUS	Sensitivitas 98,3% dan spesifisitas 100%	Meningkatkan standarisasi dan efisiensi
3	(He <i>et al.</i> 2025)	IL-24	Tidak disebutkan	Citra kanker paru	RMD-Net (deep learning)	Akurasi 95,94%	Meningkatkan interpretabilitas
4	(Wang <i>et al.</i> 2025)	P40, Pan-CK, dll	Tidak disebutkan	Kanker gastrointestinal	AI-IHC deep learning	AUC 0,90–0,96; akurasi hingga 90,81%	Kesesuaian tinggi dengan patologi
5	(Cheng <i>et al.</i> 2025)	Marker imun	Tidak disebutkan	154 kanker kolorektal	Multi-regional IHC scoring	Akurasi hingga 97,9%	Meningkatkan prediksi prognosis
6	(Krishnamurthy <i>et al.</i> 2024)	HER2	4B5, HercepTest, EP3	120 slide kanker payudara	CNN AI scoring	Akurasi 92,1%	Meningkatkan kesepakatan patologi
7	(Dunenova <i>et al.</i> 2024)	HER2	Beragam	25 studi	AI & DIA	Sensitivitas 0,94; spesifisitas 0,92	Perlu standarisasi
8	(Wen <i>et al.</i> 2024)	PD-L1, Ki-67, p53	22C3, MIB-1, DO-7	1002 slide	ResNet DL	Korelasi >0,95	Reproduksibilitas tinggi
9	(Neagu <i>et al.</i> 2024)	Multi-marker	Tidak disebutkan	Dataset retrospektif	XGBoost ML	Presisi 85,97%	Mendukung diagnosis
10	(Liao <i>et al.</i> 2024)	HER2	Tidak disebutkan	109 kasus	MembraneQuant	$\kappa = 0,77$	Mengurangi subjektivitas
11	(Kildal <i>et al.</i> 2024)	Multi-marker	Tidak disebutkan	Multi kanker	Deep learning	Akurasi hingga 91,5%	Generalisasi lintas tumor
12	(Catteau <i>et al.</i> 2023)	Ki-67	MIB-1	151 biopsi	QuPath	Kesesuaian tinggi jika ROI ditentukan	Perlu supervisi patologi
13	(Niyas <i>et al.</i> 2023)	ER, PR, HER2, Ki-67	Tidak disebutkan	600 slide	LadderNet DL	Akurasi 93,3%	Potensi otomatis penuh
14	(Mahmoud <i>et al.</i> 2023)	PRMT6	Tidak disebutkan	Kanker paru	ML scoring	Akurasi tinggi	Skalabel
15	(Zhang <i>et al.</i> 2023)	P53	Tidak disebutkan	Kanker kolorektal	P53Net DL	Akurasi 94,21%	Konsistensi tinggi
16	(Ushakov <i>et al.</i> 2023)	H-score	Tidak disebutkan	1.780 data	EndoNet DL	mAP 0,77	Lebih baik dari QuPath
17	(Xue <i>et al.</i> 2022)	Protein IHC	Tidak disebutkan	6.509 citra	ML + DNN	Akurasi 73,7%	Meningkatkan deteksi

							biomarker
18	(Elie <i>et al.</i> 2022)	Biomarker protein	Tidak disebutkan	25 kasus	GMM	Korelasi tinggi	Mengurangi bias
19	(Bencze <i>et al.</i> 2021)	Multi protein	Tidak disebutkan	Slide IHC	Pathronus AI	Lebih konsisten	Mengurangi variasi manusia
20	(Tewary & Mukhopadhyay 2021)	HER2	HercepTest	200 citra	AutoIHC	Akurasi 92,5%	Objektif dan cepat
21	(Moratin <i>et al.</i> 2021)	PD-L1, EGFR, dll	Tidak disebutkan	222 kasus	QuPath	Korelasi tinggi	Prediksi prognosis
22	(Casiraghi <i>et al.</i> 2018)	CD68, CD163	KP1, 10D6	20 pasien	ML segmentation	Akurasi >90%	Reproduksibel
23	(Jakobsen <i>et al.</i> 2018)	HER2	4B5	109 kasus	DIA	$\kappa = 0,78$	Mengurangi kasus ambigu
24	(Shu <i>et al.</i> 2016)	HER2, Ki-67	MIB-1	100 sampel	Model warna statistik	$r > 0,9$	Transparan
25	(Rizzardi <i>et al.</i> 2016)	ER β 2	Poliklonal	Prostat	DIA	Reproduksibilitas 99%	Lebih presisi
26	(Wanram <i>et al.</i> 2016)	MHC I, Tapasin	Tidak disebutkan	96 kasus	Automated IHC	κ hingga 0,755	Konsistensi meningkat

KESIMPULAN

Tinjauan sistematis terhadap 26 penelitian yang dipublikasikan pada periode 2015–2025 menunjukkan bahwa sistem kuantifikasi imunohistokimia (IHC) berbasis kecerdasan buatan secara konsisten meningkatkan objektivitas, akurasi, dan reproduksibilitas dibandingkan penilaian manual konvensional. Sebagian besar penelitian melaporkan performa diagnostik yang tinggi, dengan korelasi yang kuat terhadap penilaian ahli patologi serta kemampuan mengurangi variasi antar-pengamat dan meningkatkan efisiensi analisis. Perkembangan teknologi juga menunjukkan pergeseran dari metode analisis citra konvensional menuju pendekatan *deep learning* yang memberikan kemampuan lebih baik dalam segmentasi jaringan dan kuantifikasi biomarker pada berbagai jenis tumor. Meskipun demikian, implementasi klinis masih menghadapi tantangan berupa variasi kualitas preparat, heterogenitas jaringan, perbedaan protokol pewarnaan, serta belum adanya standar validasi dan kalibrasi yang seragam antar laboratorium. Oleh karena itu, diperlukan validasi multisenter, harmonisasi protokol, dan pengembangan standar evaluasi yang terintegrasi agar sistem AI dapat diimplementasikan secara luas dan aman dalam praktik patologi diagnostik. Berdasarkan sintesis seluruh bukti yang tersedia, AI lebih tepat diposisikan sebagai alat bantu (*decision support system*) yang mendukung, bukan menggantikan, pengambilan keputusan oleh ahli patologi.

SARAN

Berdasarkan hasil tinjauan ini, disarankan agar penelitian selanjutnya lebih difokuskan pada pengembangan dan validasi sistem kuantifikasi imunohistokimia (IHC) otomatis dengan menggunakan dataset yang lebih besar dan beragam, sehingga dapat meningkatkan generalisasi model pada berbagai kondisi klinis. Selain itu, diperlukan upaya standarisasi dalam proses pewarnaan, akuisisi citra, serta metode evaluasi kinerja algoritma agar hasil yang diperoleh dapat lebih konsisten antar laboratorium. Penelitian ke depan juga diharapkan dapat mengintegrasikan sistem berbasis kecerdasan buatan secara lebih optimal dengan peran ahli patologi, sehingga tercipta kolaborasi yang saling melengkapi antara teknologi dan keahlian manusia. Di sisi lain, pengembangan regulasi dan pedoman implementasi klinis juga menjadi hal yang penting untuk mendukung adopsi sistem ini secara luas dalam praktik patologi diagnostik.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dalam penyusunan penelitian ini. Ucapan terima kasih disampaikan kepada pihak-pihak yang telah memberikan bimbingan, arahan, dan dukungan selama proses penyusunan tinjauan ini. Penulis

juga mengapresiasi para peneliti yang karyanya menjadi referensi utama dalam penelitian ini, khususnya di bidang patologi digital dan analisis biomarker berbasis kecerdasan buatan. Semoga penelitian ini dapat memberikan manfaat dan kontribusi bagi pengembangan ilmu pengetahuan di bidang tersebut.

DAFTAR PUSTAKA

- Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., James, J.A., Salto-Tellez, M. & Hamilton, P.W., 2017, 'QuPath: Open source software for digital pathology image analysis', *Scientific Reports*, 7(1), 16878.
- Bencze, J., Szarka, M., Kóti, B., Seo, W., Hortobágyi, T.G., Bencs, V., Módis, L. V. & Hortobágyi, T., 2021, 'Comparison of Semi-Quantitative Scoring and Artificial Intelligence Aided Digital Image Analysis of Chromogenic Immunohistochemistry', *Biomolecules*, 12(1), 19.
- Brattoli, B., Mostafavi, M., Lee, T., Jung, W., Ryu, J., Park, S., Park, J., Pereira, S., Shin, S., Choi, S., Kim, H., Yoo, D., Ali, S.M., Paeng, K., Ock, C.-Y., Cho, S.I. & Kim, S., 2024, 'A universal immunohistochemistry analyzer for generalizing AI-driven assessment of immunohistochemistry across immunostains and cancer types', *npj Precision Oncology*, 8(1), 277.
- Boyaci, C., Sun, W., Hartman, J. & Ács, B., 2025, 'Global scoring method of Ki67 immunohistochemistry in breast cancer demonstrates improved concordance using real-world multi-institutional data', *Breast Cancer Research*, 27(1), 159.
- Casiraghi, E., Huber, Veronica, Frasca, Marco, Cossa, Mara, Tozzi, Matteo, Rivoltini, Licia, Leone, B.E., Villa, A. & Vergani, Barbara, 2018, 'A novel computational method for automatic segmentation, quantification and comparative analysis of immunohistochemically labeled tissue sections', *BMC Bioinformatics*, 19(S10), 357.
- Catteau, X., Zindy, E., Bouri, S., Noël, J.-C., Salmon, I. & Decaestecker, C., 2023, 'Comparison Between Manual and Automated Assessment of Ki-67 in Breast Carcinoma: Test of a Simple Method in Daily Practice', *Technology in Cancer Research & Treatment*, 22.
- Cheng, J., Han, Y., Yuan, Y., Huang, S., Xiao, B., Kong, Y., Xue, W., Yuan, R., Liu, H., Lan, P., Wu, X., Qian, Y., Ni, D. & Chen, Y., 2025, 'Automated multi-regional <sc>IHC</sc> scoring enhances prognostication in colorectal cancer', *The Journal of Pathology: Clinical Research*, 11(5).
- Choi, J.E., Kim, K.-H., Lee, Y. & Kang, D.-W., 2024, 'Digital Validation in Breast Cancer Needle Biopsies: Comparison of Histological Grade and Biomarker Expression Assessment Using Conventional Light Microscopy, Whole Slide Imaging, and Digital Image Analysis', *Journal of Personalized Medicine*, 14(3), 312.
- Dunenova, G., Kalmataeva, Z., Kaidarova, D., Dauletbaev, N., Semenova, Y., Mansurova, M., Grjibovski, A., Kassymbekova, F., Sarsembayev, A., Semenov, D. & Glushkova, N., 2024, 'The Performance and Clinical Applicability of HER2 Digital Image Analysis in Breast Cancer: A Systematic Review', *Cancers*, 16(15), 2761.
- Elie, N., Giffard, F., Blanc-Fournier, C., Morice, P.-M., Brachet, P.-E., Dutoit, S., Plancoulaine, B. & Poulain, L., 2022, 'Impact of automated methods for quantitative evaluation of immunostaining: Towards digital pathology', *Frontiers in Oncology*, 12.
- He, Z., Jia, D., Shi, Y., Li, Z., Wu, N. & Zeng, F., 2025, 'RMD-Net: A Deep Learning Framework for Automated IHC Scoring of Lung Cancer IL-24', *Mathematics*, 13(3), 417.
- Hou, C., Song, X., Chen, H., Chang, C., Lu, J., Li, C., Qu, H., Guo, R., Xu, J. & Xu, L., 2025, 'A novel automated IHC staining system for quality control application in ALK immunohistochemistry testing', *Pathology and Oncology Research*, 31, 1611964.
- Jakobsen, M.R., Teerapakpinyo, C., Shuangshoti, S. & Keelawat, S., 2018, 'Comparison between digital image analysis and visual assessment of immunohistochemical HER2 expression in breast cancer', *Pathology - Research and Practice*, 214(12), 2087–2092.
- Kildal, W., Cyll, K., Kalsnes, J., Islam, R., Julbø, F.M., Pradhan, M., Ersvær, E., Shepherd, N., Vlatkovic, L., Tekpli, X., Garred, Ø., Kristensen, G.B., Askautrud, H.A., Hveem, T.S., Danielsen, H.E., Bathen, T.F., Borgen, E., Børresen-Dale, A.-L., Engebråten, O., Fritzman, B., Hartman-Johnsen, O.J., Garred, Ø., Geisler, J., Geitvik, G.A., Hofvind, S., Kåresen, R., Langerød, A., Lingjærde, O.C., Mælandsmo, G.M., Naume, B., Russnes, H.G., Sahlberg, K.K., Sauer, T.,

- Skjerven, H.K., Schlichting, E. & Sørli, T., 2024, 'Deep learning for automated scoring of immunohistochemically stained tumour tissue sections – Validation across tumour types based on patient outcomes', *Heliyon*, 10(13), e32529.
- Krishnamurthy, S., Schnitt, S.J., Vincent-Salomon, A., Canas-Marques, R., Colon, E., Kantekure, K., Maklakovski, M., Finck, W., Thomassin, J., Globerson, Y., Bien, L., Mallel, G., Grinwald, M., Linhart, C., Sandbank, J. & Vecsler, M., 2024, 'Fully Automated Artificial Intelligence Solution for Human Epidermal Growth Factor Receptor 2 Immunohistochemistry Scoring in Breast Cancer: A Multireader Study', *JCO Precision Oncology*, (8).
- Liao, C.-H.C., Bakoglu, N., Cesmechioglu, E., Hanna, M., Pareja, F., Wen, H.Y., D'Alfonso, T.M., Brogi, E., Yagi, Y. & Ross, D.S., 2024, 'Semi-automated analysis of HER2 immunohistochemistry in invasive breast carcinoma using whole slide images: utility for interpretation in clinical practice', *Pathology and Oncology Research*, 30, 1611826.
- Mahmoud, A.M., Brister, E., David, O., Valyi-Nagy, K., Sverdllov, M., Gann, P.H. & Kim, S.J., 2023, 'Machine Learning for Digital Scoring of PRMT6 in Immunohistochemical Labeled Lung Cancer', *Cancers*, 15(18), 4582.
- Moratin, J., Mock, A., Obradovic, S., Metzger, K., Flechtenmacher, C., Zaoui, K., Fröhling, S., Jäger, D., Krauss, J., Hoffmann, J., Freier, K., Horn, D., Hess, J. & Freudlsperger, C., 2021, 'Digital Pathology Scoring of Immunohistochemical Staining Reliably Identifies Prognostic Markers and Anatomical Associations in a Large Cohort of Oral Cancers', *Frontiers in Oncology*, 11(11).
- Neagu, A.I., Poalelungi, D.G., Fulga, A., Neagu, M., Fulga, I. & Nechita, A., 2024, 'Enhanced Immunohistochemistry Interpretation with a Machine Learning-Based Expert System', *Diagnostics*, 14(17), 1853.
- Niyas, S., Bygari, R., Naik, R., Viswanath, B., Ugwekar, D., Mathew, T., Kavaya, J., Kini, J.R. & Rajan, J., 2023, 'Automated Molecular Subtyping of Breast Carcinoma Using Deep Learning Techniques', *IEEE Journal of Translational Engineering in Health and Medicine*, 11, 161–169.
- Rizzardi, A.E., Zhang, X., Vogel, R.I., Kolb, S., Geybels, M.S., Leung, Y.-K., Henriksen, J.C., Ho, S.-M., Kwak, J., Stanford, J.L. & Schmechel, S.C., 2016, 'Quantitative comparison and reproducibility of pathologist scoring and digital image analysis of estrogen receptor β 2 immunohistochemistry in prostate cancer', *Diagnostic Pathology*, 11(1), 63.
- Shu, J., Dolman, G.E., Duan, J., Qiu, G. & Ilyas, M., 2016, 'Statistical colour models: an automated digital image analysis method for quantification of histological biomarkers', *BioMedical Engineering OnLine*, 15(1), 46.
- Tewary, S. & Mukhopadhyay, S., 2021, 'HER2 Molecular Marker Scoring Using Transfer Learning and Decision Level Fusion', *Journal of Digital Imaging*, 34(3), 667–677.
- Ushakov, E., Naumov, A., Fomberg, V., Vishnyakova, P., Asaturova, A., Badlaeva, A., Tregubova, A., Karpulevich, E., Sukhikh, G. & Fatkhudinov, T., 2023, 'EndoNet: A Model for the Automatic Calculation of H-Score on Histological Slides', *Informatics*, 10(4), 90.
- Wang, J., Zhang, S., Li, J., Deng, M., Zeng, Z., Dong, Z., Chen, F., Liu, W., Wu, L. & Yu, H., 2025, 'Development and clinical validation of deep learning-based immunohistochemistry prediction models for subtyping and staging of gastrointestinal cancers', *BMC Gastroenterology*, 25(1), 494.
- Wanram, Surasak, Sombatwong, Jakkra, Sakunpong, A., Prasongdee, P., Panomket, Pawana, Wongsena, P., Limpaboon, T. & Jearanaikoon, Patcharee, 2016, 'Comparison of Automated and Conventional IHC Visual Scoring Analysis for MHC Class I and Tapasin Expression in Cervical Carcinoma.', *Journal of the Medical Association of Thailand = Chotmaihet thangphaet*, 99 Suppl 1(1), S67-75.
- Wen, Z., Luo, D., Wang, S., Rong, R., Evers, B.M., Jia, L., Fang, Y., Daoud, E. V., Yang, S., Gu, Z., Arner, E.N., Lewis, C.M., Solis Soto, L.M., Fujimoto, J., Behrens, C., Wistuba, I.I., Yang, D.M., Brekken, R.A., O'Donnell, K.A., Xie, Y. & Xiao, G., 2024, 'Deep Learning-Based H-Score Quantification of Immunohistochemistry-Stained Images', *Modern Pathology*, 37(2), 100398.
- Xue, Z.-Z., Li, C., Luo, Z.-M., Wang, S.-S. & Xu, Y.-Y., 2022, 'Automated classification of protein expression levels in immunohistochemistry images to improve the detection of cancer biomarkers', *BMC Bioinformatics*, 23(1), 470.
- Horai, Y., Mizukawa, M., Nishina, H., Nishikawa, S., Ono, Y., Takemoto, K. & Baba, N., 2019, 'Quantification of histopathological findings using a novel image analysis platform', *Journal of Toxicologic Pathology*, 32(4), 319–327.

Zhang, J., Guo, Xiaolei, Cai, W., Cao, Y., Ge, W., Huang, P., Hou, D., Zheng, S. & Zhang, G., 2023, 'Automatic Scoring Method for Tumor IHC Images Based on Deep Learning and Its Application on P53 Protein', *IEEE Access*, 11, 64005–64015.