

## SOUBORNÉ SDĚLENÍ

## Využití metod umělé inteligence v patologii

*The use of artificial intelligence methods in pathology*

Vojtěch Damian, Uladzislau Palichenka, Miroslav Koblížek, Petr Škapa, Josef Zámečník

Ústav patologie a molekulární medicíny, Fakultní nemocnice v Motole, 2. lékařská fakulta Univerzity Karlovy, Praha

## SOUHRN

**Damian V, Palichenka U, Koblížek M, Škapa P, Zámečník J. Využití metod umělé inteligence v patologii**

Současná patologie prochází zásadní proměnou. Z tradičního morfologického oboru se stává obor na pomezí klasické histopatologie, molekulární biologie a bioinformatiky. Rozvoj na poli výpočetní techniky a umělé inteligence (AI) umožňuje zpracovávat nejen data, která generují molekulárně patologické analyzátoři, ale i mikroskopické obrazy získané pomocí vysokokapacitních skenerů.

Digitalizované mikroskopické preparáty, tzv. whole slide images (WSI), je tak možné dále podrobovat algoritmickému zpracování pomocí celého spektra nástrojů strojového učení, mimo jiné také hlubokého učení založeného na umělých neuronových sítích. Právě nástroje založené na hlubokém učení otevřely nové možnosti v oblasti diagnostiky. Ukazuje se, že jsou schopné detekovat nádory či metastázy v lymfatických uzlinách, rychle odhadovat jejich grade, a dokonce predikovat molekulárně patologické změny nádorových onemocnění.

Podobně jako v ostatních medicínských oborech ani v patologii není digitalizace bez svých úskalí. Výroba histologického preparátu je především ruční prací, a tak je nutné klást velký důraz na její kvalitu, na níž je pak závislé správné fungování skenerů a následně i digitálně patologických nástrojů. Mimo tyto pro patologii specifické limitace narážíme i na překážky legislativní, ekonomické a technologické, jako je archivace a zpracování ohromných objemů dat v řádu stovek až tisíců terabytů ročně.

Digitální a výpočetní patologie představují dynamicky se vyvíjející diagnostický obor s výrazným dopadem na budoucnost nádorové diagnostiky a personalizované medicíny.

**Klíčová slova:** digitální patologie, výpočetní patologie, umělá inteligence, hluboké učení

## SUMMARY

**Damian V, Palichenka U, Koblížek M, Škapa P, Zámečník J. The use of artificial intelligence methods in pathology**

Modern pathology is currently undergoing a fundamental transformation. It is evolving from a purely morphological discipline to an integrated field that combines histopathology with molecular biology and bioinformatics. Rapid advances in computational technology and artificial intelligence (AI) have enabled the processing of not only data generated by molecular pathology analyzers, but also of microscopic images, produced by high-capacity slide scanners.

These digitized microscopic slides, also known as whole slide images (WSI), can then be algorithmically processed using a wide range of machine learning tools, including deep learning, a subset of machine learning based on artificial neural networks. These deep learning-based approaches have opened new possibilities in diagnostics. They have demonstrated the ability to, among others, detect tumors and lymph node metastases, estimate tumor grade, and even predict molecular alterations.

As in other disciplines of medicine, digitization in pathology is not without its challenges. The preparation of histological slides remains largely a manual process making it essential to maintain high-quality standards, as scanners and digital pathology tools rely heavily on clear, artefact-free slides to function properly. In addition to these pathology-specific limitations, broader economic, legal, and technical challenges must also be addressed, including the storage and analysis of massive volumes of data, which may reach hundreds or even thousands of terabytes annually.

Digital and computational pathology are rapidly advancing fields with profound implications for the future of cancer diagnostics and personalized medicine.

**Key words:** digital pathology, computational pathology, artificial intelligence, deep learning

**Korespondující autor:**

prof. MUDr. Josef Zámečník, Ph.D.  
Ústav patologie a molekulární medicíny 2. LF UK a FN v Motole  
V Úvalu 84  
150 06 Praha 5  
josef.zamecnik@lfmotol.cuni.cz

Moderní patologie jako obor tkáňové a buněčné diagnostiky už dávno není pouze oborem morfologickým. Metody molekulární biologie pronikly do patologické diagnostiky již před lety a významně ji obohatily. Detekce somatických změn v genomu u nádorových onemocnění umožňuje nejen zpřesnit diagnostiku, ale především předpovědět odpověď nádoru na cílenou léčbu – tedy uplatnit principy prediktivní patologie.

Dnes tedy patologie pracuje s celou řadou různorodých dat, ať už se jedná o obrazová data z mikroskopu, nebo o data s molekulárně genetickou informací. Jak poptávka po diagnostických i prediktivních vyšetřeních roste, roste tak i objem dat, která je nutné zpracovat. Tento nárůst však není sledován úměrným zvyšováním počtu odborníků v oboru (patologů, molekulárních biologů a bioinformatiků), do hry tak vstupuje žádoucí výpomoc počítačovým zpracováním dat sofistikovanými algoritmy strojového učení.

Rozmach umělé inteligence a digitalizace, patrný napříč vědními obory, se nevyhnul ani patologii. Klasický optický mikroskop je postupně doplňován a částečně nahrazován digitalizovanými skeny mikroskopických preparátů, tzv. **WSI (whole slide images)**, což dále umožňuje automatizované zpracování obrazu.

Modely strojového učení dnes postupně zvládají v různé míře aplikovatelnosti v praxi úkony od jednoduchého počítání buněk (například v hodnocení imunohistochemických vyšetření) přes detekci a klasifikaci vybraných nádorových onemocnění až po predikci genetických alterací některých nádorů z pouhého preparátu barveného základním barvením hematoxylinem a eozinem. Aby bylo možné takové komplexní úlohy provádět, je užíváno **hlubokého učení (deep learning)**, což je metoda strojového učení založená na využití umělých neuronových sítí.

Dnešní výzkum ve výpočetní a digitální patologii se soustředí zejména na využití tzv. **základních modelů (foundation models)**, tj. rozsáhlých umělých neuronových sítí, které jsou předtrénovány na velmi velkých objemech dat. Tyto modely se pak dají využít jako základ pro další trénování pro řešení konkrétních speciálních úloh. Nejznámějším obecným příkladem jsou modely ze série GPT, které tvoří základ nástroje ChatGPT a byly trénovány na rozsáhlých textových korpusech z celého internetu a nověji (model GPT-4) i na obrazových datech.<sup>(1)</sup>

V oblasti histopatologie můžeme uvést jako příklad model UNI z laboratoře harvardského profesora Faisala Mahmooda.<sup>(2)</sup> Je jedním z několika základních modelů, které byly vyvinuty pro účely digitální patologie, a byl předtrénován na statisících WSI různých typů tkání a onemocnění.

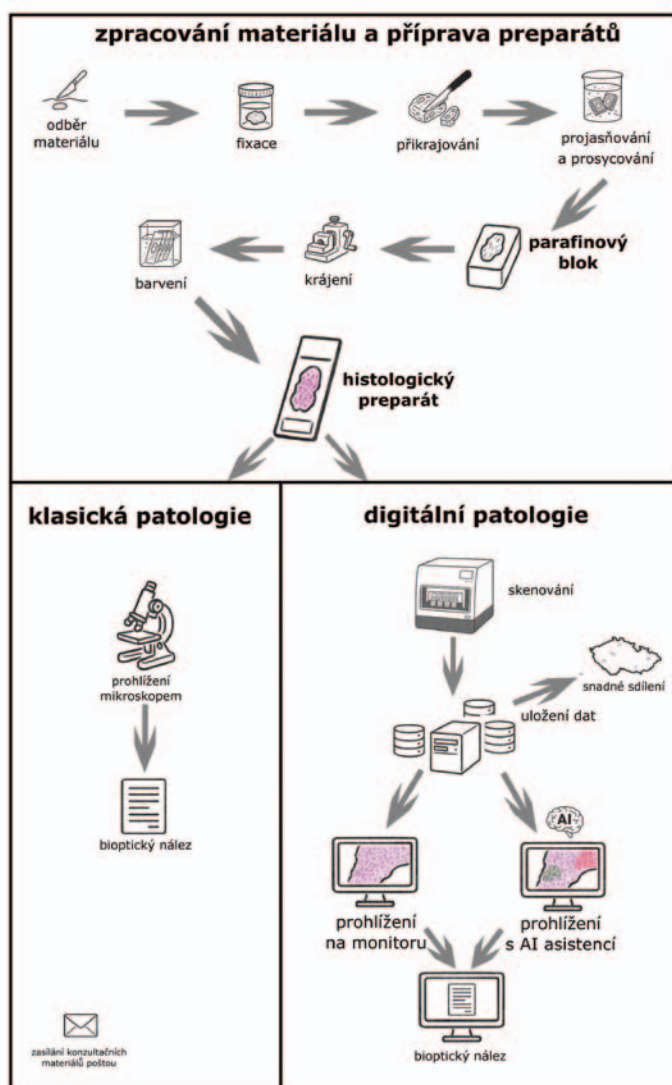
## MOLEKULÁRNÍ PATOLOGIE A ZPRACOVÁNÍ SEKVENAČNÍCH DAT

Na pracovištích patologie se dnes můžeme setkat s celou řadou analyzátorů, které generují různorodá digitální data. Větší laboratoře jsou vybaveny sekvenátory (včetně sekvenátorů nové generace – NGS) a PCR analyzátory. Objevná data (řádově desítky až stovky gigabytů, v závislosti na zkoumaném rozsahu), která jsou pomocí nich generována, je nutné jednak uchovávat, jednak dále zpracovávat a analyzovat. K archivaci molekulárních dat se užívají velkokapacitní servery, jejichž nevýhodou je prostorová náročnost a vysoká pořizovací cena, či cloudová řešení, pro něž je však nutné stabilní, kvalitní a ideálně rychlé internetové připojení. U cloudových řešení je problematická legislativní stránka, jelikož jsou citlivá data pacientů uchovávána u třetí strany mimo zdravotnické zařízení. Získaná data z molekulárně patologických vyšetření jsou pak následně zpracovávána molekulárním biologem či bioinformatikem, a to buď pomocí specializovaného komerčního softwaru, nebo pomocí lokálně vytvářených **pipelines** (tedy zjednodušeně automatizovaných postupů, které propojují různé, často volně přístupné programy, které provádějí jednotlivé úkony pro zpracování dat a pro získání výsledku).

Nástroje, které dnes bioinformatik užívá při zpracování molekulárně patologických dat, používají hlavně tradiční metody statistické analýzy a strojového učení. Hluboké učení a umělé neuronální sítě se do popředí diagnostické praxe ještě nedostaly, a to zejména kvůli vysoké přesnosti a zaužívanosti dosavadních technik. Aktuálně jsme však již ve fázi jejich počínající implementace v akreditovaných laboratořích. Hluboké učení nachází dnes uplatnění především v pokročilejších výzkumných úkolech, jako je objevování různých biomarkerů, které jsou cílitelné v rámci terapie, či mutací, které doposud nebyly asociovány s nádorovými onemocněními.<sup>(3)</sup>

## DIGITALIZACE A AI V HISTOPATOLOGII

Na rozdíl od molekulární patologie, jejíž provoz je téměř plně digitalizovaný, je histopatologie stále na počátku procesu digitalizace. I na středně velkých pracovištích patologie vznikají denně stovky analogových sklíček s barveným preparátem a k jejich převedení do digitální formy jsou potřeba velkokapacitní skenovací automaty. Takovéto skenery jsou vybaveny mikroskopem a kamerou, která postupně snímá mikroskopický obraz z celého preparátu v různé míře detailu. Výstupem jsou pak skeny WSI. V závislosti na velikosti vzorku mohou obrazové soubory dosahovat jednotek



Obr. 1: Schéma znázorňuje postup zpracování vzorků a rozdíl v práci s preparáty v rámci klasické a digitální patologie

až desítek gigabytů, a proto je i v tomto případě velmi důležitý systém ukládání a archivace dat. Obdobně jako v případě genomických dat i zde přichází v úvahu buď lokální serverové úložiště, či vzdálené cloudové řešení s výhodami a nevýhodami popsanými výše (obr. 1).

Vzniklé WSI pak může patolog sám prohlížet na monitorech s vysokým rozlišením, což mimo jiné umožňuje i vzdálené odečítání či konzultování případů (tzv. telepatologie), dále pak ale mohou být digitalizované obrazy zpracovávány pomocí modelů umělé inteligence.

Dnes jsou na trhu nižší desítky histopatologických nástrojů využívajících umělou inteligenci s certifikací IVDR, která umožňuje jejich využití v diagnostické praxi.<sup>(4)</sup> Jejich počet ale rapidně narůstá. Tyto nástroje jsou specializované na obrazy určitých typů tkání – jsou to v dnešní době zejména lymfatické uzliny, prostata a mléčná žláza. Nejčastějším úkonem, na který jsou AI nástroje designovány,

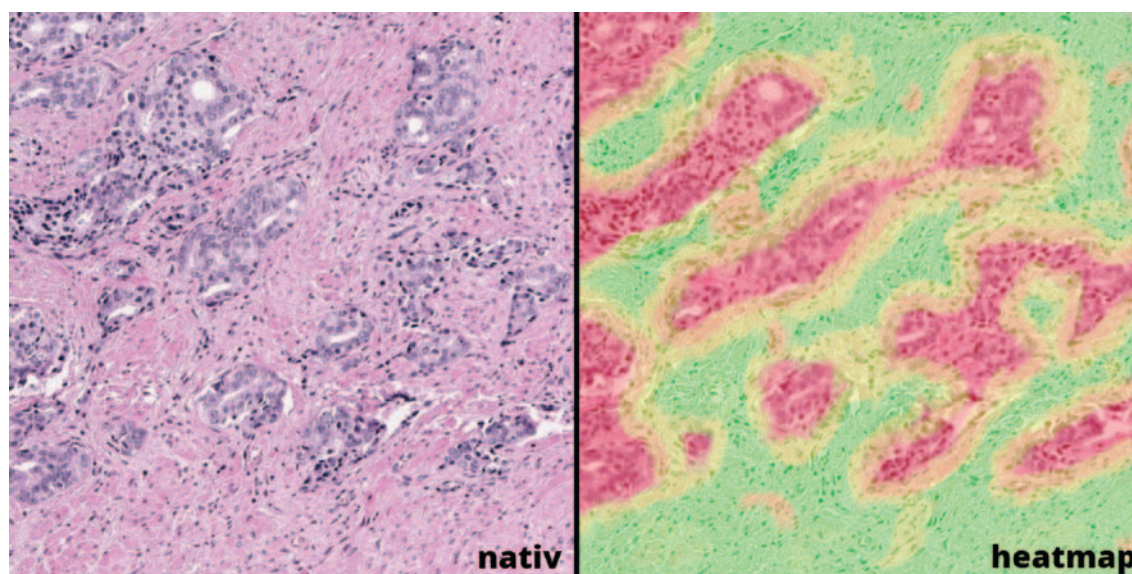
je detekce metastáz v lymfatických uzlinách nebo detekce karcinomu v jehlových biopsiích prostaty či prsu, přičemž některé nástroje nabízejí i typizaci a grading detekovaných lézí. Základní myšlenkou v těchto případech je podpořit diagnostikujícího patologa tím, že nástroj automaticky označí oblasti, kterým by měl věnovat zvýšenou pozornost. Takovou podporou se může zrychlit diagnostický proces a zároveň snížit riziko přehlédnutí drobných nenápadných ložisek. Existuje několik způsobů, kterými toho může být docíleno. Jedním z nich je rozřazení WSI na ty bez patologického nálezu a ty se suspektním ložiskem. Dalším způsobem je tvorba tzv. **heatmap** (obr. 2), což jsou vlastně dvourozměrné grafy superponované na původní WSI. Principem je, že se nad histopatologický snímek zobrazí poloprůhledná barevná vrstva rozdělující snímek na „teplejší“ a na „chladnější“ části, které v případě AI nástrojů většinou odpovídají míře pravděpodobnosti, se kterou se v daném úseku nachází suspektní léze. Oba výše zmíněné způsoby se mohou navzájem kombinovat.

## SOUČASNÝ VÝZKUM A VÝVOJ V DIGITÁLNÍ PATOLOGII

První aplikace hlubokého učení v digitální patologii se zaměřovaly na „tradiční“ úkoly, jako je grading, staging a typizace nádorů – jedna z prvních studií, která využila hluboké učení v histopatologii, byla publikována v roce 2015 a zabývala se gradingem mozkových gliomů.<sup>(5)</sup> V roce 2019 profesor Jakob Nikolas Kather ukázal, že pomocí hlubokého učení lze z WSI předpovědět stav mikrosatelitové nestability v kolorektálním karcinomu.<sup>(6)</sup> Tím položil základy nové větve výzkumu, jejímž cílem je odhalovat molekulární alterace z pouhé morfologie tkáně, tedy bez použití imunohistochemie či molekulárního testování. Navazující studie ukázaly, že stejným přístupem lze predikovat molekulární podtypy endometriálního karcinomu,<sup>(7)</sup> receptorový a genomický status karcinomu prsu<sup>(8)</sup> a také mutace genů *BRAF*, *TP53* a dalších v desítkách různých typů nádorů<sup>(9)</sup>. Takové modely by mohly v budoucnu výrazně urychlit diagnostiku nádorů i zpřístupnit komplexní vyšetření pracovištím, která nedisponují vybavením pro molekulární analýzu.

Rychlý rozvoj počítačového zpracování obrazu (tzv. počítačové vidění – computer vision) a akumulace velkého množství WSI vedly k vývoji víceúčelových **základních modelů**. Ty navazují na volně dostupné předtrénované modely, které se dále dotrénují na histologických snímcích. Trénování modelu umělé inteligence od základu je totiž velmi ekonomicky a výpočetně náročné, navíc pro to také není zatím dostupné dostatečné množství obrazových histopatologických dat. Mezi nejznámější příklady takových základních modelů patří UNI<sup>(2)</sup> a Virchow<sup>(10)</sup> založené na DINOv2<sup>(11)</sup>.

Nejpokročilejším patologickým nástrojem, který kombinuje UNI s velkým jazykovým modelem Llama 2<sup>(12)</sup>, je v tuto chvíli nástroj PathChat<sup>(13)</sup>. V praxi funguje podobně jako AI kopiloty, které jsou důvěrně známé hlavně programátorům: patolog nahraje snímky, položí dotaz a model navrhne diferenciálně diagnostické možnosti nebo třeba doporučí



Obr. 2: Preparát acinárního karcinomu prostaty. Vlevo – nativní snímek. Vpravo – „heatmap“ označující podezřelá místa ve stejném histopatologickém preparátu. Červená – suspektní karcinom; zelená – bez patologických změn.

molekulární vyšetření. Hraje tak roli partnera, který patologovi pomáhá se stanovením diagnózy. Na počátku roku 2025 získal PathChat DX od americké Food and Drug Administration status průlomového zdravotnického prostředku.<sup>(14)</sup> K aktuálnímu datu však zatím neprošel regulérním schvalovacím procesem 510(k).

Dalším perspektivním projektem je predikce výsledků prostorové transkriptomiky z hematoxylin-eozinových WSI. Prostorová transkriptomika je vyšetření, které kombinuje molekulární genetiku, přinášející informaci o genomickém profilu jednotlivých nádorových buněk, s histopatologickým vyšetřením *in situ*. Kombinace těchto technik umožňuje pozorovat rozdíly v molekulární stavbě nádorové tkáně v jejích různých oblastech. Jedná se o komplexní, časově, technicky a finančně náročné vyšetření. Tým čínských výzkumníků nedávno ukázal, že výsledky prostorové transkriptomiky lze pomocí AI do určité míry predikovat z pouhého preparátu barveného základním přehledným barvením.<sup>(15)</sup>

## ÚSKALÍ VYUŽITÍ AI V PATOLOGII

Ačkoli metody umělé inteligence mají potenciál usnadnit patologům diagnostický proces, jejich zavádění do praxe není bez úskalí. Jedním takovým úskalím je nutnost produkce vysoce kvalitních histopatologických preparátů. Histopatologie je stále převážně ruční laboratorní prací (krájení na mikrotomech, barvení, montování preparátů), tudíž záleží do velké míry na pečlivosti jednotlivého laboratorního pracovníka, jak bude výsledný preparát vypadat. Mikrotomové řezy mohou být různé silné, potrhane nebo se mohou zvlítnout při manipulaci, což následně znemožňuje skeneru nalézt vhodné ostrí body, a vzniklé WSI tak není možné vyhodnotit, a to zejména ne algoritmicky. Kvalita laboratorní práce je tak pro celý digitálně patologický proces kruciólní.

Problém může být i v barvení preparátů. Není totiž hematoxylin jako hematoxylin a eozin jako eozin a jejich odstíny se laboratoř od laboratoře nepatrně liší. Ale i takové malé rozdíly mohou působit potíže pro modely, které byly trénovány na jiných odstínech. Proto byla vyvinuta celá řada algoritmů, které se snaží barvy naskenovaných preparátů standardizovat a sjednocovat.

Kvalita WSI je důležitým determinujícím faktorem i pro vývoj modelů umělé inteligence. Pro svůj trénink vyžadují velké množství velmi kvalitních a dostatečně rozmanitých dat, která jsou pro výzkumníky poměrně často složitě dostupná. Volně dostupné jsou zejména menší datové soubory. Nejznámější dataset používaný ve výpočetně patologickém výzkumu, The Cancer Genome Atlas<sup>(16)</sup>, čítá necelých 36 000 WSI, které jsou ale značně nerovnoměrné v zastoupení jednotlivých typů tkání a jejich patologických změn. Oproti tomu žádoucí velké datasety, optimálně čítající stovky tisíc WSI, volně dostupné nejsou. Vývoj sofistikovaných modelů v patologii je tak soustředěn jen do velkých výzkumných středisek s přístupem k takovýmto velkým datovým souborům.

Také implementace výzkumných AI projektů do rutinní diagnostiky naráží na několik zásadních limitů. Především přesnost výše popisovaných molekulárních predikcí zatím zdaleka nedosahuje klinicky požadované senzitivity a specificity. Situaci navíc komplikuje absence validačních protokolů a legislativního rámce, jenž by stanovil minimální požadavky na přesnost a bezpečnost.

Další překážkou pro široké využití AI v patologii je v současné době vysoká pořizovací cena přístrojů (skenerů preparátů, velkokapacitních datových serverů) a softwaru, ale i integrace procesů do stávajících informačních systémů ve zdravotnických zařízeních. Po pořízení technického vybavení je pak nutné počítat s nemalými částkami za jeho údržbu. Podobně i software často vyžaduje periodické platby, v závislosti na konkrétní licenci. Některé licence například fungují na omezený počet vyšetření, jiné na omezený čas. Výrobci také často omezují

certifikaci výsledku na použití konkrétních, často dražších re-agencií (např. protilátek pro imunohistochemické vyšetření).

Výhledově je očekávatelné další prodražování s nutností využívání zdravotnických prostředků s IVDI certifikací.

## ZÁVĚR

Současná patologie prochází dynamickým vývojem. Posou-vá se z výhradně analogového, čistě morfologického oboru

do oboru propojujícího prvky molekulární biologie, bioin-formatiky a klasické histopatologie. Vývoj v digitální a vý-početní patologii přináší řadu možností jak na výzkumném poli, tak v klinické praxi.

Přestože řada nových nástrojů teprve na své zařazení do praxe čeká, trend je zde zřejmý. Digitalizovaný diagnos-tický postup přinese usnadnění řady procesů, jako je sdílení případů, archivace, možnost standardizace, a zejména prá-vě možnost využití algoritmického zpracování histopatolo-gického obrazu. |

## LITERATURA

1. OpenAI, Achiam J, Adler S, et al. GPT-4 Technical Report. 2024. doi: 10.48550/arXiv.2303.08774
2. Chen RJ, Ding T, Lu MY, et al. Towards a general-purpose foundation model for computational pathology. *Nat Med* 2024; 30(3): 850–862. doi: 10.1038/s41591-024-02857-3
3. Unger M, Kather JN. Deep learning in cancer genomics and histopathology. *Genome Med* 2024; 16(1): 44. doi: 10.1186/s13073-024-01315-6
4. Matthews GA, McGenity C, Bansal D, Treanor D. Public evidence on AI products for digital pathology. *Npj Digit Med* 2024; 7(1): 300. doi: 10.1038/s41746-024-01294-3
5. Ertosun MG, Rubin DL. Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. *AMIA Annu Symp Proc AMIA Symp* 2015: 1899–1908.
6. Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019; 25(7): 1054–1056. doi: 10.1038/s41591-019-0462-y
7. Fremont S, Andani S, Wolf JB, et al. Interpretable deep learning model to predict the molecular classification of endometrial cancer from haematoxylin and eosin-stained whole-slide images: a combined analysis of the PORTEC randomised trials and clinical cohorts. *Lancet Digit Health* 2023; 5(2): e71–e82. doi: 10.1016/S2589-7500(22)00210-2
8. Jaber MI, Song B, Taylor C, et al. A deep learning image-based intrinsic molecular subtype classifier of breast tumors reveals tumor heterogeneity that may affect survival. *Breast Cancer Res* 2020; 22(1): 12. doi: 10.1186/s13058-020-1248-3
9. Fu Y, Jung AW, Torne RV, et al. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nat Cancer* 2020; 1(8): 800–810. doi: 10.1038/s43018-020-0085-8
10. Vorontsov E, Bozkurt A, Casson A, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. *Nat Med* 2024; 30(10): 2924–2935. doi: 10.1038/s41591-024-03141-0
11. Oquab M, Darcet T, Moutakanni T, et al. DINOv2: Learning Robust Visual Features without Supervision. 2024. doi: 10.48550/arXiv.2304.07193
12. Touvron H, Martin L, Stone K, et al. Llama 2: open foundation and fine-tuned chat models. 2023. doi: 10.48550/arXiv.2307.09288
13. Lu MY, Chen B, Williamson DFK, et al. A multimodal generative AI copilot for human pathology. *Nature* 2024; 634(8033): 466–473. doi: 10.1038/s41586-024-07618-3
14. PathChat Receives FDA Breakthrough Device Designation | Modella AI. Dostupné na: <https://www.modella.ai/pathchat-fda-breakthrough-designation>
15. Zeng Y, Wei Z, Yu W, et al. Spatial transcriptomics prediction from histology jointly through Transformer and graph neural networks. *Brief Bioinform* 2022; 23(5): bbac297. doi: 10.1093/bib/bbac297
16. McLendon R, Friedman A, Bigner D, et al. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008; 455(7216): 1061–1068. doi: 10.1038/nature07385