Інше

Ілюшик Тарас Сергійович
магістр кафедри біомедичної кібернетики
Національного технічного університету України
"Київський політехнічний інститут імені Ігоря Сікорського"

КЛАСИФІКАЦІЯ ГІСТОЛОГІЧНИХ ЗОБРАЖЕНЬ РАКУ ПРОСТАТИ

Анотація. *Щорічно повідомляється більше 1 мільйона нових діагнозів, рак простати є другим за поширеністю онкологічним захворюванням серед чоловіків у всьому світі, що призводить до понад 350 000 смертей щорічно. Класичним методом діагностики та прогнозування є система оцінки Глісона. У цьому процесі патологоанатоми вручну аналізують зразки біопсії простати під мікроскопом, що займає багато часу і не виключає ризику отримання невірного результату. Для допомоги патологоанатомам були розроблені алгоритми глибокого навчання для виявлення раку. Багато сучасних моделей є згортковими нейронними мережами на основі патчів. Системи на основі патчів зазвичай вимагають детальних анотацій на рівні пікселів для ефективного навчання. Однак такі анотації рідко є доступними, на відміну від клінічних звітів патологів, які містять мітки на рівні слайдів. Таким чином, розробка алгоритмів, які не потребують ручних піксельних анотацій, але можуть використовувати лише клінічний звіт, була б значним прогресом у цій галузі. Проте ці анотації є важливими для надійної роботи систем комп'ютерної діагностики, оскільки вони дають пояснення результатів і дають впевненість патологоанатомам, що модель зосереджена на важливих ознаках зображення. Тому ми пропонуємо нову weakly-supervised модель глибокого навчання, основану на самонавчальних згорткових нейронних мережах, які використовують тільки глобальну оцінку Глісона гігапіксельних*

зображень під час навчання, щоб точно виконувати класифікацію патернів на рівні патчів. В цій статті ми пропонуємо використати підхід, який використовує самотреновані згорткові нейронні мережі на основі парадигми машинного навчання за набором зразків.

Ключові слова: *рак простати, нейронні мережі, оцінка Глісона, глибинне навчання, класифікація зображень.*

Виклад основного матеріалу. Рак передміхурової залози – одне з основних онкологічних захворювань у світі. На його частку припадає 14,5% усіх онкологічних захворювань у чоловіків [1], і, відповідно до Всесвітньої організації охорони здоров'я, щорічна захворюваність зростатиме до 1,8 млн. випадків цього десятиліття [2]. Шкала оцінки Глісона [3] є основним інструментом для діагностики та прогнозування. Ця система описує різні стадії раку на основі морфології біопсії простати. Процес оцінки складається з пошуку та класифікації ракової тканини на так звані шаблони Глісона (3, 4 або 5) на основі архітектурних моделей росту пухлини (рис. 1). Після того, як біопсії присвоєно оцінку Глісона, вона конвертується на оцінку ISUP за шкалою 1-5. Система оцінки Глісона є найважливішим прогностичним маркером для раку простати і ступінь ISUP відіграє вирішальну роль при вирішенні того, як слід лікувати пацієнта. Існує як ризик не виявлення раку при його наявності, так і великий ризик його розвитку, що призведе до непотрібного лікування. Однак система страждає від значної мінливості між спостерігачами між патологами, що обмежує її корисність для окремих пацієнтів.

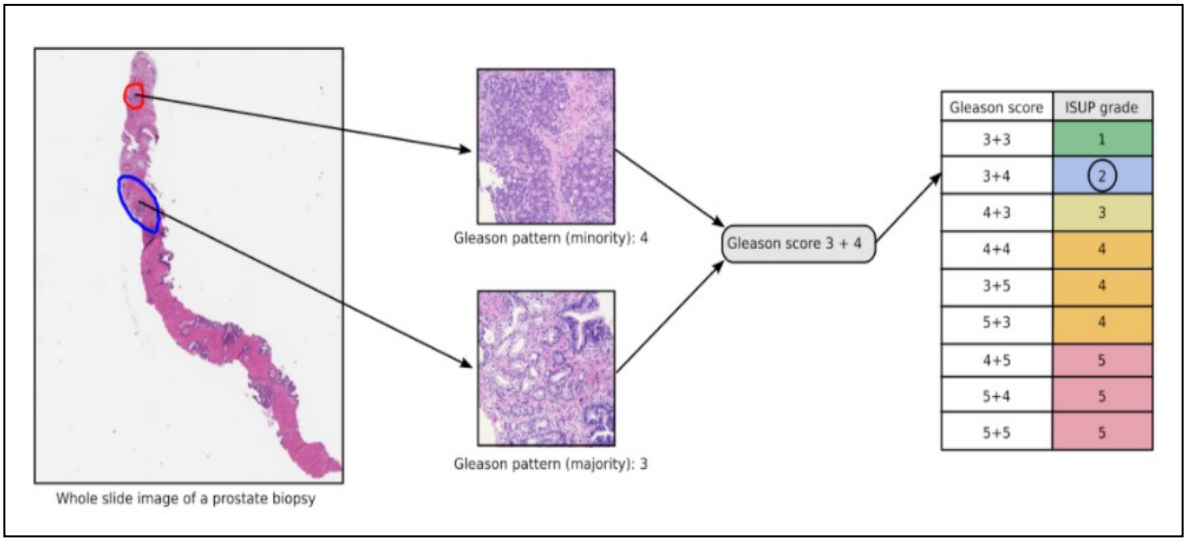

Рис 1. Ілюстрація процесу оцінювання за Глісоном біопсії, що містить рак простати

Джерело: розроблено авторами

Для того, щоб поставити діагноз раку простати, патологи отримують невеликі ділянки тканини, які обробляються за допомогою спеціального розчину. Потім за допомогою мікроскопу зразки тканини ретельно аналізуються на наявність характерних візерунків (шаблонів) відповідно до шкали Глісона. Визначаються найбільш поширені типи шаблонів. Наприклад, найбільш поширений візерунок Глісона зі значенням 3 відповідно до шкали та другий за поширеністю візерунок Глісона зі значенням 4, присутні в біопсії, визначають глобальну оцінку Глісона 3+4 = для цієї біопсії, яка, у свою чергу, перетворюється на ISUP ступені 2 згідно з рекомендаціями Міжнародного товариства урологічної патології. Якщо рак відсутній ступінь ISUP дорівнює 0.

На жаль, цей процес займає багато часу і дуже залежить від спеціаліста який проводить аналіз. Тому ці обмеження спонукають для створення автоматичних інструментів якісної і точної діагностики.

Системи автоматизованої діагностики на основі алгоритмів комп'ютерного зору здатні полегшити роботу лікарів під час діагностики. Однак є декілька факторів, які ускладнюють розробку цих систем. Як

правило, вони вимагають велику кількість даних, більш того оцифровані мікроскопічні зображення мають великий розмір і існують труднощі з отриманням анотацій на піксельному рівні [4]. Поточні системи автоматизованої діагностики зазвичай розроблені для класифікації локальні регіони раку, на основі яких вираховуються глобальна оцінка. У випадку з раком передміхурової залози потрібно вручну анотувати та розмежовувати ракові структури за допомогою системи класифікації Глісона з використанням графічних інтерфейсів користувача з різною точністю роздільної здатності. Це трудомісткий процес, схильний до помилок з боку патологів. Більш того, гетерогенний епітеліальний рак, такий як рак передміхурової залози вимагає великої кількості зразків, щоб охопити всі можливі патерни.

Література, що описує використання глобальної шкали Глісона для розробки систем комп'ютерної діагностики для класифікації біопсії простати не є такою поширеною. Основне обмеження існуючих підходів полягають у тому, що вони зосереджуються на оцінці на глобальному рівні, в той час як проблема класифікації локальних ділянок біопсії розглядається не так часто.

У цій роботі ми пропонуємо weakly-supervised навчання для виконання як глобальної оцінки біопсії, так і класифікації локальних ракових структур у тканині. По-перше, ми пропонуємо архітектуру згорткових нейронних мереж, яка здатна виявляти ракові випадки з високою достовірністю. Потім ми пропонуємо самонавчальний фреймворк, який застосовує набір даних для виконання завдання для псевдокерованого способу машинного навчання. Ми емпірично продемонструємо, що weakly-supervised моделі навчені на великих наборах даних здатні краще застосувати шкалу Глісона на рівні патчів порівняно з моделями керованого (supervised) навчання, які навчаються на менших базах даних із піксельними анотаціями. Нарешті, ми спрогнозуємо глобальну оцінку

біопсії на основі сукупності оцінок отриманих на локальних ділянках біопсії.

Аналіз останніх досліджень і публікацій.

1. Self-supervised метод

В контексті цієї роботи ми розглядали self-supervised (самонавчальний) метод машинного навчання, який має на меті використати знання про модель, яка була вперше навчена(зазвичай називається "вчителем") у другу модель (відому як "студент"). Інтерес до цієї техніки зростав останнім часом через багатообіцяючі результати, отримані під час напівконтрольованого (semi-supervised) та слабо контрольованого (weakly-supervised) сценаріїв навчання. Наприклад, у підходах напівпідконтрольного навчання, учитель використовується для отримання псевдо-міток з неанотованих даних, після цього модель тренується на анотованих даних [5; 6; 7]. Згодом, модель "студент" навчається шляхом інтеграції псевдоміток у розширений набір навчальних даних. Для навчання більш надійних моделей "учнів" вноситься шум у дані. У контексті слабокерованого навчання у деяких роботах використовується модель "вчителя" для вибору регіонів, що представляють інтерес, із зображень для навчання моделі "студент" за допомогою спрощеного набору даних [8-10].

2. Навчання за набором зразків

Метод полягає в тому, що екземпляри даних групуються в набори, які ще називаються мішками, X. Кожен мішок має мітку, Y, яка відома. У простому випадку двійкової класифікації мішок може бути позначений негативним, якщо всі екземпляри в ньому негативні. З іншого боку, мішок позначається позитивним, якщо в ньому є хоча б один екземпляр, який є позитивним.

ННЗ розглядає проблеми з неповними знаннями міток у навчальних наборах. Точніше, у навчанні за набором зразків навчальний набір

складається з міток «мішка», кожен з яких є сукупністю нерозмічених екземплярів. Мета ННЗ — передбачити розмітку для нових мішків, які раніше не зустрічались. Останні роботи з ННЗ присвячені проблемі сегментації за допомогою слабокерованого методу навчання. У цих випадках використовуються методи на основі вбудовувань (embeddings) для отримання прогнозів на рівні пікселів за допомогою градієнтних методів.

Методи. Основним методом запропонованого підходу є самоконтрольовані згорткові нейронні мережі, здатні оцінювати гістологічні зображення фрагментів передміхурової залози, використовуючи лише оцінку Глісона на рівні біопсії під час навчання.

Алгоритм поділяється на 2 моделі "вчителя" та "студента". Перша модель класифікує достовірні патчі використовуючи парадигму навчання на основі зразків (multiple instance learning) з додаванням шуму. У цьому контексті біопсія передміхурової залози розглядається як мішок X_b, що містить екземпляри $x_{b,i}$, і мета – передбачити мітки на рівні екземпляра $y_{b,i}$, коли відомі лише мітки на рівні біопсії Y_b. Y_b отримують з використанням первинної та вторинної оцінок біопсії за Глісоном, зазначених у шкалі Глісона. Потім, під час другого кроку, модель "студента" (θ_s) використовує псевдомітки на рівні екземпляра, передбачені моделлю "вчителя", для навчання на псевдонаглядовому (pseudo-supervised) наборі даних.

3. Модель "вчителя".

Модель викладача має на меті оцінити достовірні патчі за допомогою навчальних міток на рівні біопсії. Позначимо кожен окремий мішок як $X_b = \{x_{b,1}, ..., x_{b,I}\}$, де $x_{b,i}$ – i-ий екземпляр, а I позначає загальну кількість патчів, тобто екземплярів, на зображенні. Отже, метою стає передбачити глобальну оцінку Глісона з екземплярів.

Щоб зробити прогнози на рівні екземплярів, процес навчання заснований на агрегації передбачень на рівні патчів. Таким чином, для кожного екземпляра $x_{b,i}$ в сумці модель учителя прогнозує оцінку Глісона.

Потім ми використовуємо функцію агрегації, щоб відновити всі передбачення на рівні екземпляра в одне значення, яке служить висновком на глобальному рівні.

У контексті цієї роботи ми використовуємо об'єднання (pooling) як функцію агрегації. Важливо зазначити, що функція об'єднання повинна бути надійною для характеристик ННЗ. Клас на рівні мішка може бути позитивним, якщо лише один із екземплярів є позитивним для цього класу. Наприклад, використання об'єднання за середнім значенням (average pooling) призведе до зменшення глобальної активації ракових класів, якщо зображення містить велику кількість неракових патчів. Базуючись на властивостях, які спостерігаються в операції максимального об'єднання (max-pooling) в задачах сегментації із слабким контролем [11], ми пропонуємо використовувати max-pooling. Ця архітектура забезпечує класифікацію лише екземплярів з високою достовірністю, оскільки градієнти в мережі поширюються назад лише на екземплярах з найбільшою ентропією.

Для оцінки градієнту використовується бінарна крос-ентропія, як функція втрат.

4. Модель "студент".

Модель студента має на меті виконати прогнозування оцінки Глісона на рівні патчів на основі псевдоконтрольованого набору зображень, використовуючи передбачення моделі "вчителя" як псевдомітки.

По-перше, всі екземпляри з набору даних прогнозуються за допомогою моделі "вчителя". Потім виконується процес уточнення мітки на основі прогнозів моделі "вчителя" на рівні патчів та відомих глобальних міток на рівні зображень. Під час цього процесу мітки змінюються, а патчі

відкидаються. Таким чином, тільки ті патчі які отримали якусь оцінку за шкалою Глісона зберігаються для подальшого навчання моделі "студент". Що стосується неракових патчів, то їх отримують лише з відомих доброякісних предметних. Процес уточнення міток разом із спрощенням проблеми від ННЗ до псевдоконтрольованої структури дозволяє моделі "студента" краще вивчати ознаки на рівні патчів. Нарешті, модель "студента", яка має ту ж архітектуру, що й модель "вчителя", навчається, мінімізуючи категоріальну перехресну ентропію між передбаченнями та псевдомітками.

Результати

1. Дані.

Що стосується наборів даних, які використовуються для перевірки ефективності класифікаторів, набір даних MICCAI 2020 PANDA [12] використовується для оцінки запропонованих алгоритмів. Цей набір даних складається з 10 415 зображень передміхурової залози, чиї первинні та вторинні оцінки за Глісоном були позначені фахівцями-патологами. Гігапіксельні зображення були повторно дискретизовані до 10-кратної роздільної здатності та випадковим чином згруповані в три групи для навчання, перевірки та тестування. Отримані розбивки для набору даних PANDA, а також розподіл балів Глісона в обох наборах даних представлені в таблиці 1.

Таблиця 1

Розподіл даних

	NC	GS6	GS7	GS8	GS9	GS10
Train	2297	2122	2075	1002	874	99
Validation	98	89	85	42	33	6
Test	497	455	425	205	190	22
Total	2892	2666	2585	1249	1097	127

2. Метрики.

Щоб оцінити різні підходи, ми використовуємо метрику точності (ACC), оцінки f1 (F1) для класу, а також коефіцієнт каппи Коена (к). Останній метрика є основним показником якості, який зазвичай використовується при оцінці простати за Глісоном. При цьому враховується, що система Глісона складається з набору впорядкованих класів, і помилки між сусідніми класами повинні бути менш покарані.

3. Оцінка за обраними метриками.

Моделі "Вчитель Max" та "Студент Max" використовують max-pooling функцію об'єднання. Моделі "Вчитель ННЗ" та "Студент ННЗ" базуються на парадигмі навчання за набором зразків. Результати наведено в таблиці 2.

Таблиця 2

Результати досліджень

Метод	ACC	F1					k
		NC	GG3	GG4	GG5	Avg.	
Вчитель Max	0.7055	0.7263	0.7356	**0.6827**	0.666	0.7026	0.7567
Студент Max	0.722	**0.8453**	**0.7653**	0.6230	0.7029	**0.7341**	**0.7930**
Вчитель ННЗ	0.667	0.2531	0.7329	0.647	0.7253	0.5895	0.7162
Студент ННЗ	0.632	0.1342	0.7125	0.6273	**0.7333**	0.5518	0.6891

Висновки. У цій роботі ми запропонували нову методику класифікації зображень раку простати. Ми запропонували фреймворк, який поєднує в собі 2 моделі "вчителя" та "студента". Модель "студент" показала найкращий результат. Коефіцієнт Коена для цієї моделі становить 0.793 та середня оцінка f1 0.7341.

Обнадійливі результати, представлені в цій роботі, є значними прогресом у літературі з гістології простати. Користуючись weakly-supervised методом навчання нейронних мереж, можна оцінювати локальні шаблони на зображеннях перевершуючи supervised методи, які потребують

трудомісткості анотації від патологів. Подальші дослідження будуть зосереджені на вивчення та вдосконалення процесу нормалізації зображення простати гістологічні зразки для використання систем комп'ютерної діагностики.

Література

1. World Cancer Research Foundation, "Prostate cancer statistics". 2019. [Online]. URL: http://www.wcrf.org
2. World Health Organization, "Global cancer observatory". 2019. [Online]. URL: http://gco.iarc.fr
3. Gleason D. F. Histologic grading of prostate cancer: A perspective, human pathology. 1992.
4. Epstein J. I. A new contemporary prostate cancer grading system. Annales de Pathologie. 2015. Vol. 35, No. 6. PP. 474–476.
5. Burchardt M., Engers R., M¨uller M., Burchardt T., Willers R., Epstein J. I., Ackermann R., Gabbert H. E., A De La Taille., Rubin M. A. Interobserver reproducibility of Gleason grading: Evaluation using prostate cancer tissue microarrays. Journal of Cancer Research and Clinical Oncology. 2008. Vol. 134, No. 10. PP. 1071–1078.
6. Komura D., Ishikawa S. Machine Learning Methods for Histopathological Image Analysis. Computational and Structural Biotechnology Journal. 2018. Vol. 16. PP. 34–42. DOI: https://doi.org/10.1016/j.csbj.2018.01.001
7. Alldrin Veit N., Chechik G., Krasin I., Gupta A., Belongie S. Learning from noisy large-scale datasets with minimal supervision. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR. 2017. Vol. 2017-Janua. PP. 6575–6583.
8. Bazzani L., Bergamo A., Anguelov D., Torresani L. Self-taught object localization with deep networks. 2016 IEEE Winter Conference on Applications of Computer Vision. 2016.

9. Sangineto E., Nabi M., Culibrk D., Sebe N. Self Paced Deep Learning for Weakly Supervised Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019. Vol. 41, No. 3. PP. 712–725.
10. Jie Z., Wei Y., Jin X., Feng J., Liu W. Deep self-taught learning for weakly supervised object localization. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition. 2017. Vol. 2017-Janua. PP. 4294–4302.
11. Bulten W., Litjens G., Pinckaers H., Strm P., Eklund M., Kartasalo K., Demkin M., Dane S. The PANDA challenge: Prostate cANcer graDe Assessment using the Gleason grading system. Mar. 2020. DOI: https://doi.org/10.5281/zenodo.3715938