

## **ВИКОРИСТАННЯ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ АВТОМАТИЧНОЇ ДІАГНОСТИКИ РАКУ ЛЕГЕНІВ**

**Б.В. ЧАПАЛЮК, Ю.П. ЗАЙЧЕНКО**

**Анотація.** Рак легенів є одним із найбільш агресивним та швидко прогресуючим онкологічним захворюванням, що призводить до смерті великої кількості людей по всьому світу. Одним із основним методів боротьби з раком є виявлення його якомога на найбільш ранній стадії. Побудова автоматичних систем діагностування раку легенів може допомогти лікарям в цьому. Розглянуто можливість побудови рекурентної нейронної мережі для аналізу КТ зображень легенів пацієнтів. Описано побудовану нейронну мережу, що складається із згорткової нейронної мережі, рекурентної нейронної мережі та додаткового механізму уваги, що надає можливість перевикористати попередньо анотовану інформацію про злоякісні ділянки на КТ знімку.

**Ключові слова:** рекурентні нейронні мережі, глибинне навчання, механізм уваги, автоматична діагностика раку легенів

### **ВСТУП**

Розроблення систем автоматичного оброблення медичних зображень — одне з найбільш важливих завдань, вирішивши яке, можна зберегти мільйони життів у всьому світі та поліпшити медичне обслуговування. Це спонукає багатьох науковців та інженерів створювати подібні системи, сподіваючись пришвидшити та підвищити точність медичного діагностування. Мета роботи — розгляд проблеми побудови системи автоматичного діагностування раку легенів на базі рекурентних нейронних мереж.

Рак легенів є однією з найбільш поширених причин смерті пацієнтів, що мають проблеми з онкологією. Причиною цього є агресивність і складність перебігу захворювання, а також проблема виявлення захворювання на пізніх стадіях. Відповідно до даних Всесвітньої організації охорони здоров'я (World Health Organization) протягом 2012 р. зафіксовано близько 1,8 млн нових випадків захворювання раком легенів і близько 1,6 млн випадків смертей [1]. Рак легенів вважається одним із найбільш агресивних видів ракових захворювань через велику кількість летальних випадків у пацієнтів. Протягом п'яти років після діагностування раку легенів виживає лише 10–15% пацієнтів. Проте, якщо рак вдається діагностувати на ранніх стадіях, то протягом десяти років виживає близько 75% пацієнтів, а отже, діагно-

стування раку легенів на дуже ранніх стадіях є життєво необхідною умовою для збереження життя, тому людям із групи ризику призначають регулярні перевірки. Щоб допомогти лікарям проводити подібні діагностування швидко та якісно, розробляються системи автоматичного діагностування раку легенів.

Основним методом діагностування раку легенів є отримання рентгенівського знімка легені або тривимірного зображення за допомогою комп'ютерної томографії (КТ) шляхом пошарового пропускання рентгенівського променя крізь тонкі шари тканин людського тіла в різних напрямках. У результаті отримується набір двовимірних зображень частини тіла людини, з'єднання яких дає тривимірне зображення. У роботі використовувались тривимірні зображення легенів, отримані завдяки комп'ютерній томографії.

Для виявлення злоякісних утворень на знімку рентгенологи зазвичай шукають вузлові утворення діаметром понад шість міліметрів. Якщо вузлове утворення має менший діаметр, імовірність того що вузлик є злоякісною пухлиною дуже низька, але він може стати зачатком злоякісної пухлини, тому лікарі періодично перевіряють їх збільшення з часом отриманням КТ знімка легенів один раз за рік. Переглядаючи та аналізуючи КТ зображення легенів, рентгенологи ретельно перевіряють тканини шар за шаром, аналізуючи просторові залежності між підозрілими вузликами. За таким підходом зроблено спробу застосувати схожий метод аналізу раку легенів у системі автоматичної діагностики.

## ОГЛЯД ІСНУЮЧИХ РІШЕНЬ ТА ЛІТЕРАТУРИ

Проблема автоматичного діагностування раку легенів може бути розглянута як задача класифікації злоякісної пухлини на зображенні КТ, що являє собою тривимірне зображення легенів пацієнта. Зазвичай розмір КТ зображення дуже великий, тому помістити його в пам'ять відеокарти разом з більш-менш великою нейронною мережею неможливо. Наприклад, набір даних DSB 2017 [2] має розмір  $512 \times 512 \times (\text{кількість шарів})$ . Кількість шарів у зображенні КТ залежить від роздільної здатності сканера, за допомогою якого виконується обстеження пацієнта. У деяких випадках набір даних містить близько 350 шарів. Оскільки вимоги до обчислювальної потужності пам'яті машини підвищуються кубічно з розміром тривимірного зображення, існує декілька підходів для побудови систем діагностування з урахуванням таких обмежень.

Кожен шар КТ зображення можна опрацьовувати як окремо, так і з серією двовимірних даних, використовуючи вже відомі архітектури двовимірних згорткових нейронних мереж — DenseNet [3], ResNet [4] або Xception [5]. Система будується таким чином: кожен шар КТ зображення пропускається через двовимірну згорткову нейронну мережу, яка на виході повертає ймовірність на шарі злоякісної пухлини. Після отримання результатів для кожного із зображень рішення про наявність приймається на основі припущення задачі багатозразкового навчання (Multi-instance learning); якщо хоча б один шар КТ зображення містить злоякісну пухлину, то вважається що все КТ зображення містить злоякісну пухлину. Якщо на жодному шарі зображення не виявлено злоякісної пухлини, то вважається, що все КТ зображення не містить злоякісного утворення.

Однією з найбільших переваг використання такого підходу є можливість перевикористання великої кількості навчених на ImageNet згорткових нейронних мереж. Не зважаючи на те, що вони навчені на даних, які не належать до знімків легенів або медичної галузі, виявлені низькорівневі ознаки на одному наборі даних є корисними і пришвидшують навчання нейронної мережі на іншому наборі даних, а також підвищують точність роботи нейронної мережі [6]. Однак такий підхід ігнорує тривимірну природу даних, а отже, ігнорує просторові залежності між шарами знімка КТ. Тому такий підхід працює, але його можливості досить обмежені.

Іншим підходом є створення тривимірних згорткових нейронних мереж, що можуть використовувати тривимірні дані на вході. Для перетворення КТ зображення у тривимірне потрібно об'єднати всі шари зображення, наклавши двовимірні зображення один на один. Тоді таке зображення можна подавати на вхід тривимірної нейронної мережі. Як згорткову нейронну мережу можна використовувати такі моделі, як C3D [7], або 3D DenseNet [8]. Але зазвичай в медичному домені розмічених даних небагато, і їх не достатньо для того, щоб навчити такі нейронні мережі повністю. Тому підходи [9] передбачають два етапи: сегментацію та класифікацію. На етапі сегментації тривимірне згорткове нейронна мережа намагається знайти ділянки, на яких можливе злоякісне утворення. Для навчання цієї нейронної мережі набір даних має містити інформацію про ділянки, які вказують на наявність потенційно злоякісних утворень. Таку інформацію попередньо описують рентгенологи. Після етапу сегментації виявлені підозрілі ділянки легень подають на тривимірну згорткову нейронну мережу, що прогнозує ймовірність раку в кожній з підозрілих ділянок. Якщо хоч в одній з вибраних ділянок ймовірність раку буде високою, система розглядатиме все тривимірне зображення як таке, що містить рак легенів. Це збігається із припущенням у багатозразковому навчанні. Як показують праці [9, 10], тривимірні згорткові нейронні мережі можуть досягати точності роботи рівня людини, що становить близько 87% правильних відповідей. Система може також надати інформацію про ділянки з виявленими пухлинами, що спрощує взаємодію лікарів з подібними автоматизованими системами.

Попри всі досягнення і переваги використання тривимірних згорткових нейронних мереж, зроблено спробу побудувати систему, яка працює за принципом того, як це робить людина, і перевірити, яких результатів можна досягти з таким підходом. Професійний рентгенолог, досліджуючи КТ знімка пацієнта, переглядає кожне зображення легенів, виявляючи ділянки, які містять потенційно злоякісні утворення. Перевіряючи кожний шар, рентгенолог запам'ятовує інформацію про підозрілі ділянки на попередньому шарі і намагається пов'язати цю інформацію з тим, що є на поточному шарі, враховуючи просторові залежності, розміщення, форму та морфологію. Подібний підхід можна спробувати реалізувати в автоматизованих системах за допомогою рекурентних нейронних мереж.

## **ОПИС ВИКОРИСТАНИХ НАБОРІВ ДАНИХ**

Для діагностування та перевірки наявності раку легенів зазвичай використовують КТ. Зображення, отримане за допомогою КТ, складається із серії рентгенівських знімків, узятих під різним кутом з різних положень. Такі знімки

можуть містити зображення різних тканин, кісток та кровоносних судин. У випадку знімка КТ для діагностування раку легенів зображення робитиметься на рівні грудної клітки пацієнта і міститиме зображення легеневої тканини. Якщо об'єднати отримані зображення в одне, наклавши їх одне на одне, матимемо тривимірне зображення легенів, яке надалі можна використовувати як вхід для нейронної мережі. Діагностуючи рак легенів, рентгенологи шукають ділянки, на яких є злоякісні скупчення, що утворюють вузлики. Зазвичай під час діагностування намагаються віднайти всі вузлики, навіть якщо вони не дуже великі. Небезпечними і найбільш імовірними злоякісними утвореннями вважаються вузлики розміром 6 мм і більше. Малі вузлики перевіряють протягом певного періоду часу на наявність їх росту. Якщо малий вузлик за півроку чи рік значно збільшився, це може свідчити про рак легенів.

Експерименти та навчання нейронної мережі проводилися з використання наборів даних LUNA [11] і Data Science Bowl 2017 (DSB 2017) [2]. DSB 2017 містить КТ зображення для більше ніж 1000 пацієнтів, кожен з яких містить інформацію про наявність раку легенів на знімку. Інформацію про наявність раку підтверджено діагнозом в лікарні. DSB 2017 розбито на тренувальну, перевірку та тестову вибірки, які містять 1397, 198 та 506 зображень КТ пацієнтів відповідно. Тренувальна вибірка є не збалансованою і має 1035 прикладів без наявних ознак раку та 362 зображення-прикладі, що містять пухлину. Набір даних LUNA включає 1186 міток ділянок, що позначають потенційно злоякісну пухлину у 888 пацієнтів. Дані були отримані й описані професійними рентгенологами. Приклади потенційно злоякісних утворень часто використовуються для вирішення завдань сегментації. У випадку побудованої рекурентної нейронної мережі в роботі використовували ановані ділянки в механізмі уваги.

## ПОПЕРЕДНЄ ОБРОБЛЕННЯ ДАНИХ

Дані в КТ знімках містяться в шкалі одиниць Гаунсфілда, що кількісно відображає густину рентгенівського випромінювання. Усі КТ сканери налаштовані та відкалібровані точно вимірювати ці значення. За замовчуванням у наборі даних DSB 2017 наявні не в шкалі Гаунсфілда, тому перед початком використання їх потрібно конвертувати. Це можна зробити шляхом перемноження значень пікселів зображення та параметра «rescale slope» і значень параметра «rescale intercept» [12]. Ці параметри задаються виробником сканера і доступні в метаданих dicom файлу. У наборі даних LUNA це робити не потрібно, оскільки тут значення пікселів уже містяться в шкалі Гаунсфілда.

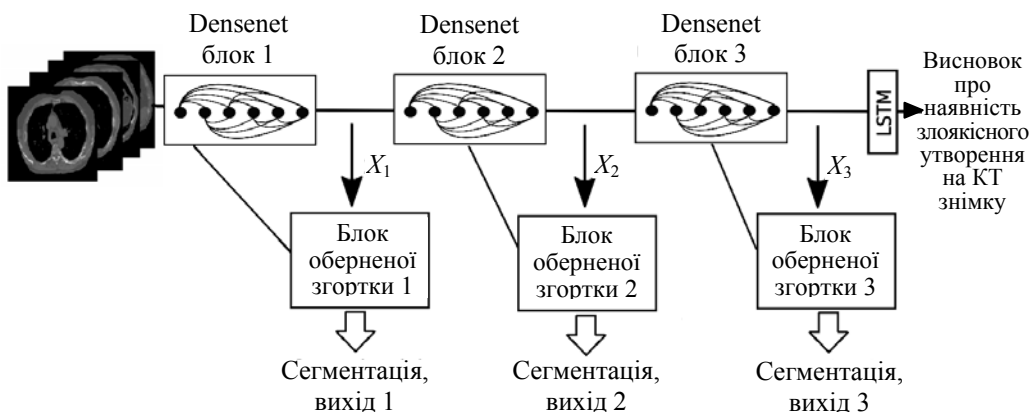
У шкалі Гаунсфілда кожний вид біологічної тканини, кістки, кров, судини мають свої значення рентгенівської густини. Цей факт широко використовують для відфільтрування із зображення, що не стосується легенів. Для цього було відфільтровано всі значення пікселів, які більші або дорівнювали 600. Останнім кроком у попередньому обробленні є зміна розміру зображення та нормалізація значень пікселів для пришвидшення навчання нейронної мережі. У розгляданому випадку кожен шар КТ зображення зменшено до розміру 200×200 для зменшення розмірності вхідних даних.

## ПОБУДОВА ТА НАВЧАННЯ РЕКУРЕНТНОЇ НЕЙРОННОЇ МЕРЕЖІ

У дослідженні вирішено спробувати реалізувати рекурентну нейронну мережу, яка б розглядала КТ знімок у спосіб, подібний до того, як це робить людина. Для цього потрібно поєднати між собою двовимірну згорткову нейронну мережу та рекурентну нейронну мережу. Як двовимірну згорткову нейронну мережу використано DenseNet [3], тоді як рекурентна нейронна мережа реалізована у вигляді LSTM [13]. У цьому випадку згорткова нейронна мережа є екстрактором ознак, який виділяє найбільш важливі ознаки в кожному шарі КТ знімка, а LSTM прогнозує наявність злоякісного утворення в легенях, аналізуючи послідовність отриманих ознак із згорткової нейронної мережі. Тобто рекурентна мережа навчається аналізувати інформацію і залежності між різними шарами КТ знімка з урахуванням тривимірної природи. Спробою навчити таку модель отримано точність роботи нейронної мережі на рівні 0,72, що значно нижче від точності, досягнутої тривимірними згортковими нейронними мережами. Основною проблемою навчання такої мережі є недостатня кількість прикладів у тренувальному наборі даних, що зумовило серйозні проблеми перенавчання.

Для поліпшення роботи нейронної мережі потрібно було розділити задачу на підзадачі, як і у випадку тривимірних згорткових нейронних мереж. У праці [14] автори будували систему для розпізнавання пошкоджень у головному мозку людини на базі КТ знімків черепа людини. Основна особливість побудованої нейронної мережі полягає у використанні механізму уваги [15], яка надавала інформацію нейронній мережі про локацію злоякісного утворення на знімку під час етапу навчання. Локація злоякісного утворення бралася із набору даних, попередньо описаного професійними рентгенологами. Такий підхід значно підвищував точність роботи нейронної мережі зі знаходження пошкоджень у головному мозку людини. Тому було вирішено використати схожий механізм для підвищення точності роботи рекурентної нейронної мережі для виконання поставленого завдання.

В існуючу модель DenseNet + LSTM після кожного Densenet блока додали блоки оберненої згортки (див. рисунок).



Кінцева версія рекурентної нейронної мережі для прогнозування раку на КТ знімку

Додані блоки оберненої згортки під час навчання використовуються як механізм додавання інформації про ділянки злоякісного утворення і виконують сегментацію. Такий механізм надає нейронній мережі інформацію

про місце виявлення проблемної ділянки та стимулює більше приділяти увагу саме потенційно небезпечним ділянкам. Блок оберненої згортки складається з декількох послідовно з'єднаних шарів згортки  $1 \times 1$  та оберненої згортки для відновлення вихідного розміру зображення. Між відповідним DenseNet блоком та шарами з блока оберненої згортки існують додаткові зв'язки, зроблені в такій ж манері, як і в нейронній мережі типу U-Net [16]. Це необхідно для покращення роботи шару оберненої згортки.

Для навчання створеної рекурентної нейронної мережі з механізмом уваги необхідно виконати додаткові кроки оброблення тренувальної вибірки. У наборі даних LUNA [11] є попередні описані рентгенологами координати знаходження потенційно злоякісних утворень, з використанням цієї інформації створено маски, які застосовуються для порівняння того, що повертає нейронна мережа після кожного з блоків оберненої згортки і включення цієї інформації у функцію втрат. Набір даних DSB 2017 не містить інформації про злоякісні утворення в легенях, тож перевикористано описи підозрілих ділянок, створених в межах праці [9].

Під час навчання нейронної мережі використовували DenseNet з глибиною 121, яка була попередньо навчена на наборі даних ImageNet. Використання попередньо навченої нейронної мережі дозволило отримати кращу початкову ініціалізацію вагових коефіцієнтів мережі, що пришвидшило процес навчання. Фінальну функцію втрат визначено як суму з ваговими коефіцієнтами між функціями втрат для задачі класифікації та задач сегментації (визначених блоками оберненої згортки). Навчання нейронної мережі відбувалося за допомогою стохастичного градієнтного спуску з початковим значенням коефіцієнта навчання (learning rate) 0,0005. Упродовж навчання коефіцієнт навчання змінювався вручну залежно від динаміки навчання побудованої мережі. У результаті навчена нейронна мережа показала точність роботи на рівні 0,81.

## РЕЗУЛЬТАТИ

У роботі були навчені два різні види рекурентних нейронних мереж, які описано в попередньому розділі. Перша рекурентна нейронна мережа складалася з двовимірної згорткової нейронної мережі DenseNet [3] та LSTM [13]. Інша модель додає до DenseNet та LSTM механізм уваги шляхом додавання трьох додаткових задач, що розв'язують задачу сегментації (див. рисунок). Також додано значення точності тривимірних нейронних мереж, отримані іншими авторами, для порівняння результатів, здобутих для рекурентної нейронної мережі з іншими підходами.

Результати проведених експериментів та порівняння з тривимірними згортковими нейронними мережами наведено в таблиці.

Models	Accuracy
DenseNet + LSTM	0,72
DenseNet + LSTM з механізмом уваги	0,81
DSB 2017 winner model [9]	0,87

## ВИСНОВКИ

Побудована рекурентна нейронна мережа показала задовільні результати, що не дуже відрізняються від рівня точності людини, але все ще гірші за результати, отримані за допомогою тривимірних згорткових нейронних мереж. Результати показують, що для кращого навчання і роботи рекурентної нейронної мережі необхідно використовувати механізм уваги, що допомагає нейронній мережі навчитися звертати увагу на дійсно злякисні утворення. Сам механізм уваги реалізовується як додаткові задачі сегментації, які дають змогу враховувати інформацію про злякисні ділянки, що були попередньо ановані рентгенологами. Основною перевагою використаного механізму уваги є можливість перевикористання людського знання для покращення і пришвидшення навчання мережі.

## ЛІТЕРАТУРА

1. World Health Organization, «World cancer report». — 2014. — Available at: <http://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-2014>.
2. Kaggle, Data Science Bowl. — 2017. — Available at: <https://www.kaggle.com/c/data-science-bowl-2017>.
3. Huang G. Densely Connected Convolutional Networks / G. Huang, Z. Liu, L. van der Maaten, K. Q Weinberger // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — P. 2261–2269. — Available at: <https://www.doi.org/10.1109/CVPR.2017.243>
4. He K. Deep Residual Learning for Image Recognition / K. He, X. Zhang, S. Ren, J. Sun // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2016. — P. 770–778. — Available at: <https://www.doi.org/10.1109/CVPR.2017.243>
5. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions / F. Chollet // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — P. 1800–1807. — Available at: <https://www.doi.org/10.1109/CVPR.2017.195>
6. Pan S. A Survey on Transfer Learning / S. Pan, Q. Yang // IEEE Transactions on Knowledge and Data Engineering. — 2010. — Vol. 22, N. 10. — Available at: <https://www.doi.org/1345-1359.10.1109/TKDE.2009.191>
7. Tran Du. Learning Spatiotemporal Features with 3D Convolutional Networks / Du. Tran, L. Bourdev, R. Fergus et al. // Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). — 2015. — P. 4489–4497. — Available at: <https://www.doi.org/10.1109/ICCV.2015.510>
8. Bui T.D. 3D Densely Convolutional Networks for Volumetric Segmentation / T.D. Bui, J. Shin, T. Moon. — 2017. — Available at: <https://arxiv.org/abs/1709.03199>.
9. Liao F. Evaluate the Malignancy of Pulmonary Nodules Using the 3D Deep Leaky Noisy-or Network / F. Liao, M. Liang, Z. Li et al. // IEEE transactions on neural networks and learning systems. — 2017. — Available at: <https://www.doi.org/10.1109/TNNLS.2019.2892409>
10. Zhu W. DeepLung: Deep 3D dual path nets for automated pulmonary nodule detection and classification / W. Zhu, C. Liu, W. Fan, X. Xie // Proceedings — 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018. —

- Vol. 2018-Janua. — 2018. — P. 673–681. — Available at: <https://www.doi.org/10.1109/WACV.2018.00079>
11. *Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge // Medical Image Analysis.* — 2017. — Vol. 42. — P. 1–13. — Available at: <https://www.doi.org/10.1016/j.media.2017.06.015>
  12. *DicomLookup.* — Available at: <http://dicomlookup.com/lookup.asp?sw=Tnumber&q=%280028,1052%29>
  13. *Greff K. LSTM: A Search Space Odyssey / K. Greff, R. Srivastava, J. Koutník et al. // IEEE Transactions on Neural Networks and Learning Systems.* — Vol. 28, N 10. — 2017. — P. 2222–2232. — Available at: <https://www.doi.org/10.1109/TNNLS.2016.2582924>
  14. *Grewal M. RADNET: Radiologist Level Accuracy using Deep Learning for HEMORRHAGE detection in CT Scans / M. Grewal, M.M. Srivastava, K. Pulkit, V. Srikrishna.* — 2017. — Available at: <https://arxiv.org/abs/1710.04934>
  15. *Mnih V. Recurrent Models of Visual Attention / V. Mnih, N. Heess, A. Graves, K. Kavukcuoglu.* — Available at: <https://arxiv.org/abs/1406.6247>
  16. *Ronneberger O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science.* — Springer, 2015. — P. 234–241. — Available at: [https://www.doi.org/10.1007/978-3-319-24574-4\\_28](https://www.doi.org/10.1007/978-3-319-24574-4_28)

Надійшла 16.09.2019