

АНАЛІЗ ФОТОЕМІСІЙНИХ СПЕКТРІВ БАГАТОЗОННИХ НАДПРОВІДНИКІВ МЕТОДАМИ МАШИННОГО НАВЧАННЯ

К. Г. Богачов^{1, а}, О. А. Кордюк^{2,3, б}

¹Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»,
Фізико-технічний інститут

²Інститут металофізики імені Г.В. Курдюмова НАН України

³Київський академічний університет

Анотація

Запропоновано спосіб розділення електронних зон у фото-емісійних спектрах шляхом використання згорткових нейронних мереж. Метод застосовано до проблеми виявлення "двошарового розщеплення" у надпровідних купратах, зокрема у Bi-2212. Показано процес генерації даних для цих надпровідників. Проведено аналіз результатів на здатність узагальнення, стійкість до шуму, та апробовано на реальних спектрах.

Ключові слова: ARPES, High-temperature superconductors, Convolutional Neural Networks

Вступ

Основна задача використання машинного навчання для проблеми надпровідності формулюється як прогнозування матеріалів за заданими параметрами. Жодна з теорій надпровідності не дає рецепт, що треба зробити, щоб підняти критичну температуру. Всі пошуки нових високо-температурних надпровідників зводяться до майже випадкового пошуку. Використовуючи машинне навчання, ми можемо спробувати розв'язати цю задачу. Треба навчити машину які матеріали та сполуки мають яку критичну температуру й, відповідно, поставити задачу пошуку високотемпературного надпровідника. Або існує схожа задача навчити машину розуміти природу надпровідників. Ці задачі можна вирішувати через електронну структуру.

Електронна зонна структура — це розподіл електронів в імпульсно-енергетичному просторі. Вона, разом з поверхнею Фермі, були важливими теоретичними концептами, на яких ґрунтувалось багато теорій фізичних властивостей твердих речовин. Наразі, завдяки розробці відповідних експериментальних методів, ці концепції перевірено безпосередньо спостережуваними проявами квантової природи електронів у кристалах. Серед нових експериментальних методів цікавою є кутова фото-емісійна спектроскопія з кутовим розділенням — ARPES [1].

Можна сказати, що ARPES дозволяє бачити реальний розподіл електронів в імпульсно-енергетичному просторі кристала та допомагає зрозуміти процеси розсіювання та взаємодії електронів з іншими складовими кристалу.

1. Методологія дослідження

Для дослідження властивостей багатозонних надпровідних купратів використовувались згорткові нейронні мережі. Але для їх навчання необхідно мати фотоемісійні спектри. Приклад зображень наведено на рисунку 1. Проблема полягає у тому, що необхідно доволі велика кількість (10000 – 100000) таких зображень, запис яких потребував би занадто багато часу. Однак можна піти іншим шляхом та згенерувати спектри.

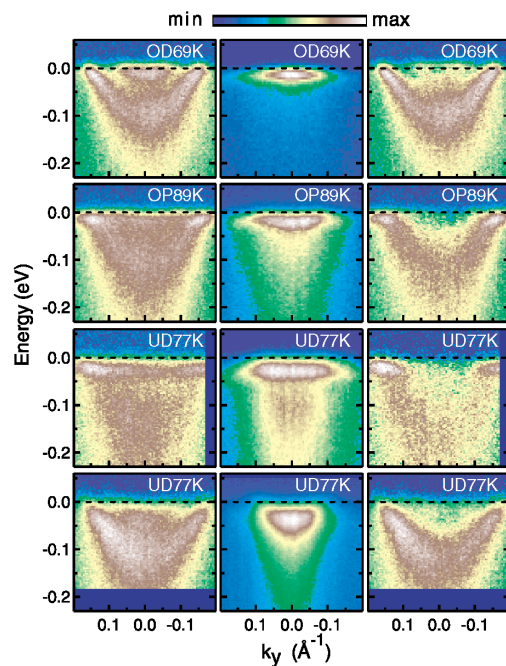


Рис. 1. Приклад ARPES спектрів як функцій енергії та хвильового вектора вздовж напрямку $(\pi - \pi) - (\pi, -\pi)$ [2]

^аbogachov.kyrylo@gmail.com

^бkordyuk@gmail.com

Генерація датасету

Спектральна функція являє собою уявну частину функції Гріна одноелектронних збуджень квазічастинок [1].

$$A(w, \mathbf{k}) = -\frac{1}{\pi} * \text{Im}(\mathbb{G}) \quad (1)$$

Враховуючи взаємодію, функцію Гріна можна записати:

$$\mathbb{G} = \frac{1}{\omega - \varepsilon - \Sigma}, \quad (2)$$

$$\Sigma = \Sigma' + i\Sigma'' \quad (3)$$

Тоді одночастинкова спектральна функція:

$$A(w, \mathbf{k}) = \frac{1}{\pi} * \frac{\Sigma''(\omega)}{(\omega - \varepsilon(\mathbf{k}) - \Sigma'(\omega))^2 + \Sigma''(\omega)^2} \quad (4)$$

ARPES забезпечує пряме вимірювання фотоструму $I(\omega, \mathbf{k}, T)$. Власне повну картину можна отримати тільки враховуючи функцію розподілу Фермі-Дірака,

$$f(\omega, T) = \frac{1}{1 + e^{\frac{E - E_f}{k_B T}}} \quad (5)$$

матричні коефіцієнти та вклад декількох зон.

Варто відзначити, що на початку 2000 – x , у період активного розвитку ARPES-експерименту (див. посилання в [1]), вважалося, що не дивлячись на наявність двох CuO шарів у кристалічній будові "двошарових" купратів, таких як Bi-2212 ($\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+x}$), відповідне енергетичне розщеплення відсутнє у надпровідному стані [3]. Саме це прогнозувала теорія лауреата Нобелівської премії Філіпа Андерсона, яка передбачала "ув'язнення" в CuO шарі [4]. Цікаво, що не дивлячись на експериментальні спростування ефекту "ув'язнення" [2], наявність двошарового розщеплення нижче T_c для купратів з низькою концентрацією носіїв є і досі дискусійною темою. Це пов'язано з суттєвим розмиттям спектрів за рахунок сильної залежності $\Sigma(\omega)$, яке є до того ж суттєво немонотонним завдяки взаємодії з "магнітним резонансом" [5]. Вирішення дискусійного питання наявності двошарового розщеплення методом машинного навчання і є темою даної роботи. Для генерації датасету використовувалась наступна модель, що передбачала можливість існування двох зон:

$$I(\omega, \mathbf{k}, T) = M_a(h\nu, \mathbf{n}, k) * f(\omega, T) * A(w, \mathbf{k}) + M_b(h\nu, \mathbf{n}, k) * f(\omega, T) * A(w, \mathbf{k}) \quad (6)$$

Було обрано випадок аномально сильної залежності власної енергії від її уявної частини. Уявну частину власної енергії та дисперсію невзаємодіючих електронів можна представити наступним чином:

$$\Sigma''(w, \mathbf{k}) = C + \alpha w^2 + \beta T^2 + \frac{G^2}{(w - w_0)^2 + G^2}, \quad (7)$$

$$\varepsilon(\mathbf{k}) = \varepsilon_0 - t \cos(\mathbf{k}) \quad (8)$$

Згорткова нейронна мережа

Розмір вхідного зображення було змінено на 224×224 , а самі зображення віднормовані по 3 каналам RGB $[0, 255] \rightarrow [0, 1]$. Архітектура моделі представлена на рисунку 2. В якості основного екстрактора інформації з зображення використовувалась предтренована на Imagenet згорткова мережа ResNet50 (Residual Network). Мережа ResNet була проривним кроком в покращенні результатів різних задач комп'ютерного зору, адже вона дозволяє збільшувати кількість шарів, при цьому збільшуючи точність. Реалізовано це додатковим зв'язком між собою двох послідовних блоків мережі, що дозволяє зменшити проблему зникаючого градієнту [6].

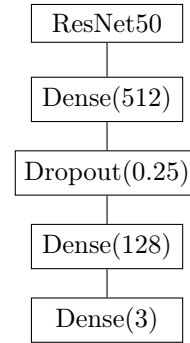


Рис. 2. Архітектура моделі

Після усереднювального агрегувального шару (average pooling) ResNet повертає матрицю розмірності $[batchsize, 2048]$, де $batchsize = 128$ — розмір однієї підгрупи вхідних даних. Наступні шари $Dense(n)$ — плоскі шари, призначені для обрахунку параметрів задачі, n — кількість нейронів, $Dropout(p)$ — шар призначений для регуляризації. При кожній наступній ітерації тренування випадково обрані вузли виключаються з мережі з імовірністю p . Це дозволяє більш рівномірно збалансувати ваги між нейронами [7].

В якості функції втрат використовувався метод найменших квадратів. Призначення функцій втрат полягає в обчисленні кількості характеристики, яку модель повинна прагнути мінімізувати під час навчання. Для цієї мінімізації було використано метод Adam з параметрами $(\beta_1 = 0.9, \beta_2 = 0.999, \alpha = 0.0001)$ [8]. На виході маємо три параметри M_a, M_b, dw_0 .

2. Результати

Підставляючи вирази (4, 5, 7, 8) в (6) було згенеровано 10000 спектрів. Для оцінки якості тренування датасет був розділений на 9000 зображень в навчальній і 1000 зображень у валідаційній вибірці. Декілька таких зображень наведено на рисунку 3. Тренування тривало 50 епох та близько 2 годин на GPU. Результати, точність та функції втрат останньої епохи для тренувальної й валідаційної вибірки наведені у таблиці 1.

Порівнюючи точність на тренувальному наборі та валідаційному приходимо до висновку, що модель не перетренована, оскільки значення до-

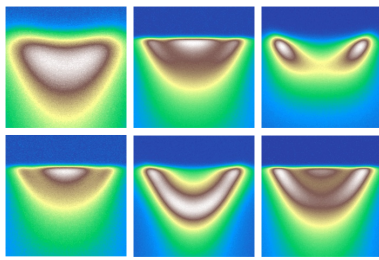


Рис. 3. Приклад згенерованих ARPES спектрів

волі схожі. Це також означає, що модель добре вміє узагальнювати, тобто не намагається повторювати результати тренувальної вибірки. Параметри

M_a, M_b, dw_0 дають повну інформацію про структуру спектру та наявність в ній кількох зон. На рисунку 4 наведено результати на реальних спектрах. З експерименту відомо, що сплітінг двох ліній цього надпровідника dw_0 має складати близько 0.1eV. Також було проведено аналіз залежності

Табл. 1. Результати тренування

Train accuracy	93.94%
Train loss	0.0024
Validation accuracy	91.74%
Validation loss	0.0053

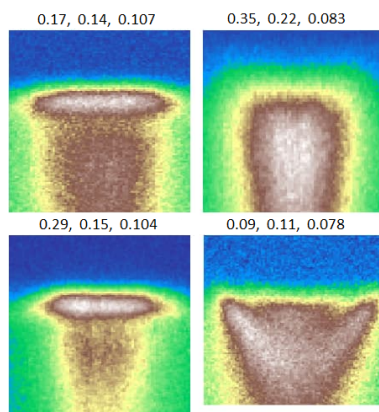
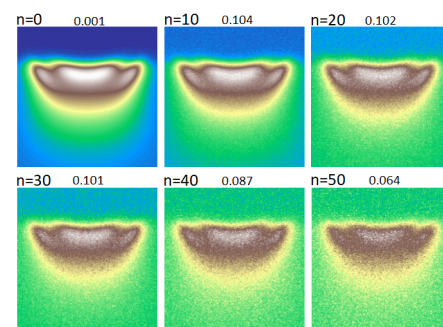


Рис. 4. Результати на реальних спектрах, по порядку M_a, M_b, dw_0 .

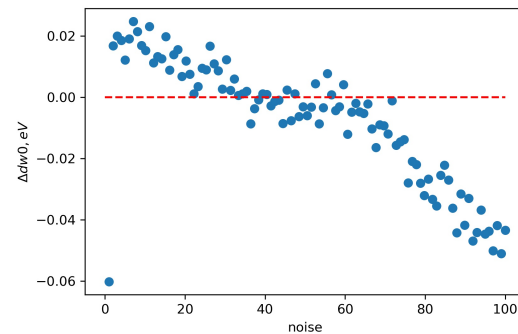
точності визначення параметрів як функції шуму. На рисунку 5а наведено візуалізацію одного й того самого спектру зі сплітінгом $dw_0 = 0.1\text{eV}$. На рисунку 5б наведено графік залежності похибки визначення dw_0 від шуму. Значення бралось не абсолютне, тобто $\Delta dw_0 = \text{prediction} - \text{truth}$. Видно сильний викид за відсутності шуму, що пояснюється тим, що модель тренувалась за наявності шуму на кожному зображенні, а також наявна спадна залежність.

Висновки

У роботі було створено згорткову нейронну мережу для аналізу ARPES спектрів надпровідників. Система була натренована з використанням згенерованих даних та протестована на реальних спектрах. Фінальна точність передбачень склала 92%. Також було проведено аналіз якості передбачень мережі



(а) Спектри з різним рівнем шуму. Число зліва — рівень шуму, по центру — dw_0



(б) Графік залежності похибки визначення dw_0 від шуму

Рис. 5. Залежності dw_0 від рівня шуму

як функції шуму. Використовуючи такий самий метод, можна діставати параметри з більш складних моделей або генерувати власні.

Роботу було виконано у рамках проекту НФДУ 2020.02/0408 "Багатозонність електронних станів: фізика та застосування".

Перелік використаних джерел

1. Kordyuk A. A. ARPES experiment in fermiology of quasi-2D metals // Low Temperature Physics. — 2014. — Vol. 40. — P. 375–388.
2. Kim T. K., et al. // Phys. Rev. Lett. 91.167002. — 2003.
3. Ding H., et al. // Phys. Rev. Lett. 76, 1533. — 1996.
4. Anderson P. W. Experimental Constraints on the Theory of High-Tc Superconductivity // Science. — 1992. — Vol. 256.
5. Dahm T., et al. // Nature Physics. — 2009. — Vol. 5. — P. 217–221.
6. He Kaiming, et al. Deep Residual Learning for Image Recognition. — 2015. — Access mode: <https://arxiv.org/abs/1512.03385>.
7. Improving neural networks by preventing co-adaptation of feature detectors / Hinton G. E., Srivastava N., Krizhevsky A., Sutskever I., and Salakhutdinov R. R. — 2012. — Access mode: <https://arxiv.org/abs/1207.0580>.
8. Kingma Diederik P., Lei Jimmy. ADAM: A Method For Stochastic Optimization. — 2015. — Access mode: <https://arxiv.org/abs/1412.6980>.