

Ю.В. ПУСТОВІТ, Є.П. ЛИТВЕНЮК, Є.Д. ЛІМАРЕВ

Київський національний університет імені Тараса Шевченка

(Вул. Володимирська, 64/13, Київ 01601; e-mail: jura.pustvit@gmail.com)

МЕТОД ОБРОБКИ ARPES СПЕКТРІВ НА ОСНОВІ U-NET

УДК 539

Фотоемісійна спектроскопія з розділенням по куту (ARPES) – це потужний інструмент для дослідження електронної структури матеріалів. Однак визначення електронної дисперсії з цих спектрів може бути ускладненим через ефекти уширення, наявність різних типів шуму, обмеження роздільної здатності тощо. У цій статті пропонується новий підхід до визначення дисперсії зі спектрів ARPES на основі нейронної мережі U-Net. Задача виділення енергетичних зон розглядається як задача семантичної сегментації. Ми показуємо, що U-Net, навчена лише на основі згенерованих даних, може визначити зонну структуру з експериментально отриманих спектрів без попереднього знешумлення.

Ключові слова: ARPES, нейронні мережі, електронна структура, U-Net.

1. Вступ

Важливість машинного навчання та нейронних мереж у фізиці стрімко зростає. Ці технології доводять свою здатність вирішувати складні проблеми, які раніше вважалися нерозв'язними. Галузі, де вони мають значний вплив, досить широкі: від астрономії [1] до фізики високих енергій [2]. Проблеми, що вирішуються за допомогою цих підходів, також різночоті відрізняються: від простого знешумлювання зображень до ідентифікації бозона Хіггса [1, 2].

Фотоемісійна спектроскопія з розділенням по куту (angle-resolved photoemission spectroscopy, ARPES) є однією з експериментальних методик, яка в останні роки отримала багато вдосконалень, таких як ARPES з часовим розрізненням, зі спіновим розрізненням, з просторовим розрізненням тощо [3, 4]. Однак класичні методи, такі як EDC та MDC аналізи [5, 6] або різні методи обробки зображень (друга похідна [7], кривина [8], мінімальний градієнт [9]), не підходять для роботи з великими обсягами даних зі складною розмірністю. Тому сьогодні для обробки складніших експерименталь-

них даних запропоновано різні методи машинного навчання. Такі методи дають можливість знешумлювати зображення за допомогою різних підходів (на основі згорткової нейронної мережі [10] або методів, що не потребують тренувального набору [11, 12]), вирізняти (візуалізувати) особливості спектрів [13], або виділяти ознаки та знешумлювати зображення за допомогою автокодувальника [14].

У цій статті запропоновано новий підхід до визначення дисперсії електронних зон з експериментально отриманих ARPES спектрів. Проблема визначення електронної дисперсії розглядається як задача семантичної сегментації. Мета полягає в тому, щоб класифікувати кожен піксель зображення до певного класу. Запропонований метод використовує нейронну мережу U-Net з блоком уваги для розрізнення пікселів “фону” та електронної дисперсії. Як було показано, така мережа може обробляти експериментально отримані спектри без попередньої обробки, навіть якщо для тренування та валідації використовуються лише згенеровані дані.

2. Метод

Метод ARPES дозволяє пряме вимірювання електронної зонної структури [4]. Однак ця структура може зазнавати уширення та спотворення через вплив різних факторів, включаючи електрон-електронну та електрон-фононну взаємодії, а також кристалічні домішки. Ці фактори можуть спричиняти зсуви енергії та розширення піків.

Цитування: Пустовіт Ю.В., Литвенюк Є.П., Лімарев Є.Д. Метод обробки ARPES спектрів на основі U-Net. *Укр. фіз. журн.* **70**, № 8, 557 (2025).

© Видавець ВД “Академперіодика” НАН України, 2025. Стаття опублікована за умовами відкритого доступу за ліцензією CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

ISSN 2071-0194. *Укр. фіз. журн.* 2025. Т. 70, № 8

Спектри також зазнають впливу шумів (дробовий, фон Ширлі) та обмеження роздільної здатності. Для отримання неспотвореної електронної дисперсії використовується кілька методів. Вони включають такі аналітичні методи як MDC та EDC, та різні методи обробки зображень, такі як методи другої похідної, кривини та мінімального градієнта.

Нещодавно для визначення електронної дисперсії були запропоновані різні методи на основі нейронних мереж [13, 14]. Наприклад, визначення дисперсії електронної зони з зашумлених та ушпирених спектрів здійснюється шляхом розв'язання задачі оберненого перетворення [13]. Ідея полягає у використанні згорткової нейронної мережі (ЗНМ) для знаходження відповідного оберненого перетворення для спектра. Хоча запропонований метод дозволяє отримувати кращі результати, ніж вищенаведені традиційні методи, для нього є характерними певні недоліки. По-перше, зображення має бути попередньо оброблене (знешумлене) для визначення дисперсії, тому деякі деталі можуть бути втрачені. По-друге, ЗНМ враховує зв'язки між сусідніми пікселями, але не враховує цієї залежності на великих відстанях. Обмежене рецептивне поле ЗНМ (на кожен піксель впливає лише його околиця 13×13 пікселів) може не охоплювати всі важливі зв'язки між віддаленими пікселями, які впливають на електронну дисперсію [13].

Наш підхід розглядає визначення електронної дисперсії як задачу сегментації. Спектр розглядається як карта сегментації, де кожен піксель класифікується в один із двох класів. Перший клас – це пікселі, що відносяться до електронної дисперсії. Другий клас охоплює фонові пікселі, пов'язані з ушпиренням, шумом та іншими артефактами. Ця задача схожа на задачу видалення фону. У нашому випадку “фон” включає всі пікселі, які не пов'язані безпосередньо з електронною дисперсією.

2.1. U-Net з механізмом “уваги”

Одним з підходів, розроблених на основі нейронних мереж, що найчастіше використовується для семантичної сегментації, є нейронна мережа U-Net та її варіації [15]. U-Net спочатку була запропонована для сегментації біомедичних зображень. Для візуалізації положень електронних енергетичних зон ми використали модифіковану мережу U-Net з механізмом “уваги” (Attention-U-Net) [16]. Ця ме-

режа містить додатковий блок уваги, який захоплює достатньо велике рецептивне поле, що дозволяє враховувати залежності на великих відстанях без значного ускладнення моделі U-Net [16].

Архітектура U-Net складається з трьох основних частин: кодувальника (канал стиснення), симетричного декодувальника (канал ушпирення) та “прихованого” шару між ними. Основний будівельний блок цієї нейронної мережі включає повторне (двократне) застосування двох 3×3 -згорток, після застосування яких не змінюється розмірність зображення та кожна з яких супроводжується пакетною нормалізацією та ReLU активацією.

Кодувальник складається з чотирьох блоків, за кожним з яких розташований агрегувальний шар максимального узгодження для зменшення розмірності. Кількість ядер у згорткових шарах відповідає шаблону $64 \times n$, де n відповідає номеру блоку.

Латентний шар застосовує два згорткові про-шарки (вони не змінюють розмірність зображення) з 64×16 ядрами, за кожним з яких відбувається пакетна нормалізація та ReLU активація.

Декодувальник відтворює кодувальник з чотирма блоками. На кожному кроці декодувальника блок уваги обробляє сигнал з відповідного блоку кодувальника.

Вихідні дані після блоку уваги конкатенуються з ознаками з попереднього блоку декодера зі збільшеною роздільною здатністю, використовуючи блок апсемплінгу (блок підвищення роздільної здатності до 2×2), та обробляються блоком декодера. Кількість ядер у згорткових шарах відповідає шаблону $64 \times 8/2^{n-1}$, де n дорівнює номеру блоку декодувальника.

Кінцевий шар 1×1 – це згортка з двома ядрами, після якого проводиться пакетна нормалізація та сигмоїдна активація. Він використовується для відображення виходу декодувальника на карту сегментації класів.

Пікселі зображення, що пов'язані з дисперсією електронних зон, займають лише невелику його частину, тому існує значний дисбаланс між пікселями, які належать до різних класів. Для вирішення проблеми дисбалансу класів була використана фокальна втрата (focal loss) як функція втрат [17]. Її головною перевагою є поділ зразків на легкі та складні. Легкий зразок вже добре класифікований (ймовірність віднесення до класу близька до

одиниці), тому його внесок можна значно зменшити або домножити на певний ваговий коефіцієнт. У цьому випадку внески легких зразків у загальну функцію втрат невеликі, навіть якщо їх кількість значна. Модель зосереджена на тренуванні підмножини складних зразків, для яких передбачувана ймовірність значно менша за одиницю [17]. Це дозволяє уникнути ситуацій, коли модель досягає майже 100% точності, просто відносячи все до класу, що частіше зустрічається на зображенні.

2.2. Процес тренування

Основною перешкодою в процесі тренування є відсутність доступних тренувальних даних, отриманих з експериментальних спектрів. Це зумовлено складнощами процесу визначення дисперсії, оскільки цей процес складно автоматизувати, а у випадку визначення “на око” результат буде не достатньо точним [13]. Тому було використано згенерований набір даних, де положення електронної дисперсії може бути точно визначено [10,13,14]. Як правило, набір, який використовується для тренування нейронної мережі для обробки спектрів ARPES, досить малий (менше ніж 10000 зразків) [10,13,14].

Згенерований набір даних було використано для тренування запропонованої нейронної мережі U-Net. Навчальний набір даних нейронної мережі складається з пар зображень: на вході – це згенеровані спектри, які можуть містити до двох зон, а на виході – це відповідне промічене зображення (карта сегментації), де кожен піксель зображення має певну мітку. Розмірність вхідного зображення становить $128 \times 128 \times 1$, а вихідного – $128 \times 128 \times 2$. Кожен піксель вихідного сигналу міститься за допомогою унітарного кодування вектора: $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ або $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (піксель дисперсії електронної зони). Використовується бінарний вектор, де лише один елемент дорівнює 1, показуючи чи піксель належить до електронної дисперсії чи фону. Для моделювання експериментально отриманих однозонних спектрів використовувалась одночастинкова спектральна функція, а саме уявна частина функції Гріна для одноелектронних збуджень (квазічастинок),

$$A(\omega, k) = \frac{\Sigma''(\omega)}{[\omega - \varepsilon(k) - \Sigma'(\omega)]^2 + [\Sigma''(\omega)]^2},$$

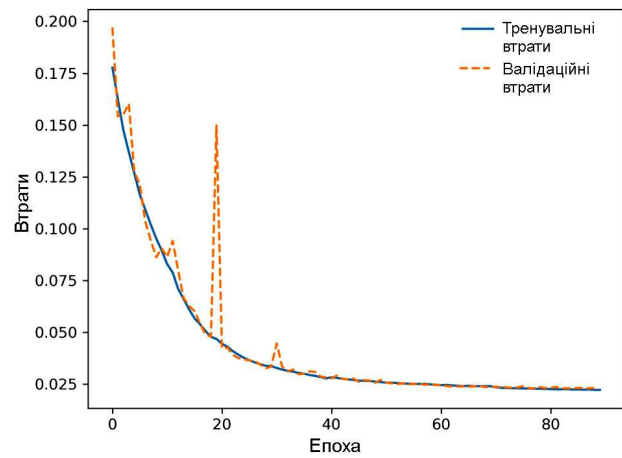


Рис. 1. Функція втрат протягом процесу тренування

Значення параметрів

Параметр	Мін.	Макс.
α	0	1,5
λ	2,1	8
l	-0,3	1,3
m	1	20
Imp	0	1

де $\Sigma'(\omega)$ та $\Sigma''(\omega)$ – це дійсна та уявна частини власної енергії квазічастинки, відповідно, що відображає всю взаємодію електронів у кристалі; $\varepsilon(k) = mk^2 + l$ – це “гола” дисперсія електронів (апроксимується простою параболою); а $\Sigma'(\omega) = \lambda\omega$ та $\Sigma''^2 = \text{Imp}$, де постійний параметр Imp додається для врахування розсіювання на домішках. Значення параметрів були випадковим чином згенеровані з діапазону значень, представлених у таблиці.

Оскільки роздільна здатність спектра зазвичай перевищує 128×128 пікселів, при обробці він розділявся на менші підобласті, які можна аналізувати. Порівняно з усім спектром, відносно малий розмір цих підобластей означає низьку ймовірність того, що будь-яка задана підобласть містить більше двох електронних зон. Тому ми використали тренувальні дані, які включали спектри, що містять до двох зон. Двостонні спектри були одержані шляхом додавання двох однозонних спектрів.

Оскільки експериментальні дані завжди містять шум, до змодельованих спектрів було додано два типи шуму, а саме:

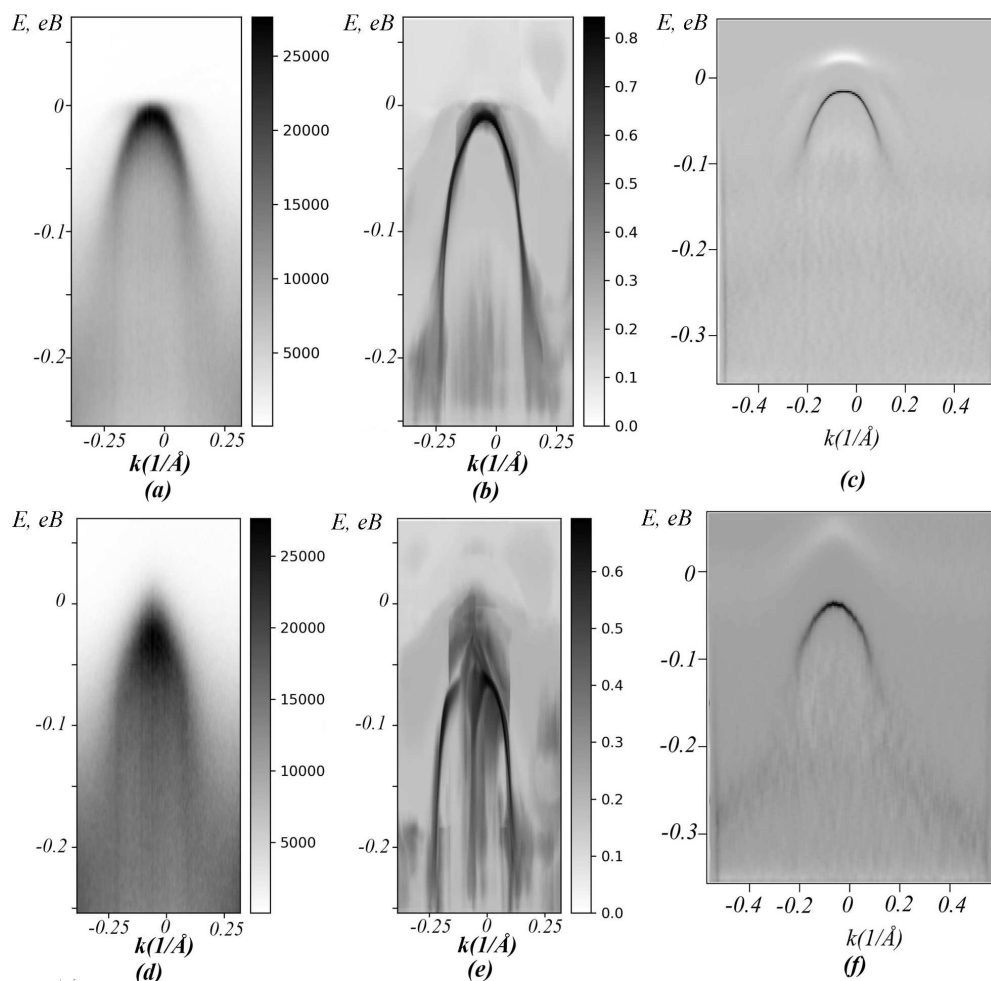


Рис. 2. Спектри ARPES у точці Z зони Брілюєна в Fe(Se, Te), отримані з використанням лінійної вертикальної поляризації для температур 20 К (а) та 220 К (d), а також відповідні результати їх обробки за допомогою нашого методу U-Net (b), (e) та методу кривини (c), (f)

- шум Ширлі для опису фону, пов'язаного з непружно розсіяними електронами ($\beta\omega^2$, де β має значення в інтервалі від 0 до 1,5);

- шум Пуассона – кожен згенерований піксель зображення множиться на випадкове значення з рівномірним розподілом (від 0 до 0,5) та додається до зображення.

Згенеровані спектри домножувалися на розподіл Фермі–Дірака з температурами в діапазоні від 0 до 300 К для моделювання ефектів, пов'язаних з рівнем Фермі.

Щоб зробити процес тренування нейронної мережі менш трудомістким та для економії пам'яті, процес тренування складався з кількох фаз.

1. Початкове тренування (20 епох). У початковій фазі тренування використовувався набір даних з 1000 зразків (тренувальний набір – 900 зразків, та валідаційний набір – 100 зразків).

2. У другій фазі відбувалося розширення набору даних. Початкові 1000 зразків замінювалися новим набором даних із 1500 зразків (1350 для тренування та 150 для валідації).

3. Новий набір даних складався ще з 1500 зразків (1350 для тренування та 150 для валідації) та використовувався протягом 30 епох.

4. Нарешті проводилося остаточне розширення набору даних та завершення тренування (20 епох). Тривалість тренування скорочувалася до 20 епох,

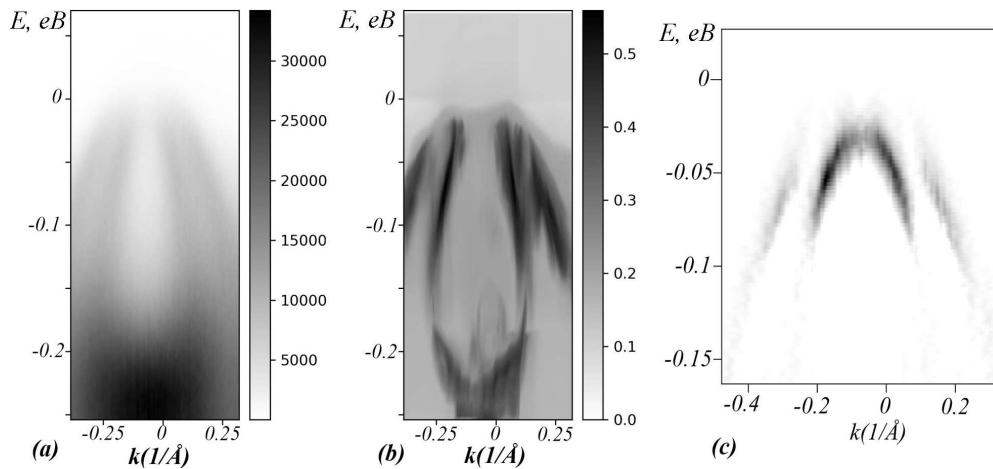


Рис. 3. Спектр ARPES у точці Z зони Брілюєна в $\text{Fe}(\text{Se}, \text{Te})$, отриманий з використанням лінійної горизонтальної поляризації для температури 160 K (a) та відповідні результати її обробки за допомогою нашого методу U-Net (b) та методу кривини (c)

а набір даних розширювався до 2000 зразків (1800 для тренування та 200 для валідації).

Такі розширення та зміни набору даних надають нейронній мережі ширший спектр зразків, покращують її здатність до узагальнення, та запобігають її перетренуванню. Це гарантує, що мережа матиме достатньо часу для тренування за новим набором даних та підтримки своєї загальної продуктивності.

Оптимізатор Adam використовувався для тренування мережі протягом 90 епох. На рис. 1 показано, що значення функції втрат для тренувального та валідаційного наборів монотонно зменшуються та стає майже постійним протягом останніх 10 епох, що свідчить про добру збіжність. Зокрема, значення функцій втрат для обох наборів дуже схожі.

3. Результати

Для порівняння продуктивності та надійності методу його було застосовано до експериментально отриманих спектрів у точці Z зони Брілюєна в $\text{Fe}(\text{Se}, \text{Te})$. Спектри були отримані за різних температур (від 20 до 220 K), що дозволило відстежувати продуктивність мережі за різних рівнів шуму. Крім того, спектри охоплюють як одно-, так і двозонні конфігурації. Оскільки розмір зображення перевищує 128×128 пікселів, передбачення моделі були розраховані для всіх можливих ділянок зображення розміром 128×128 . Піксель з найвищим

значенням можливої ймовірності того, що він належить до електронної дисперсії, був обраний для ділянок, що перекриваються. Цей підхід допомагає врахувати всі можливі залежності на великих відстанях між пікселями зображення, доступними при такій роздільній здатності нейронної мережі.

Ми порівняли результати нашого методу з результатами, отриманими за допомогою методу кривини в роботі [18], щоб оцінити його продуктивність. Однак у роботі [18] кожен спектр був попередньо оброблений шляхом згладжування для зменшення шуму перед застосуванням методу кривини. Результати методу кривини були оптимізовані шляхом підбору параметра C_0 для досягнення найкращого результату [8]. Як показано на рис. 2, наш метод дозволяє візуалізувати дисперсію в ширшому діапазоні енергій, демонструючи більшу стійкість до шуму, що вноситься непружно розсіяними електронами. Враховуючи залежності пікселів на великих відстанях, мережа може вловити зоноподібну природу електронної дисперсії. Як показано на рис. 2, e, мережа уширює картину дисперсії, отриману на основі раніше передбачених пікселів. На відміну від методу кривини та більшості інших методів, наш підхід нейронної мережі безпосередньо обробляє експериментальні спектри без необхідності їх попереднього згладжування.

Здатність нейронних мереж оминати процедуру згладжування є важливою при роботі зі спектрами, що містять кілька близько розташованих

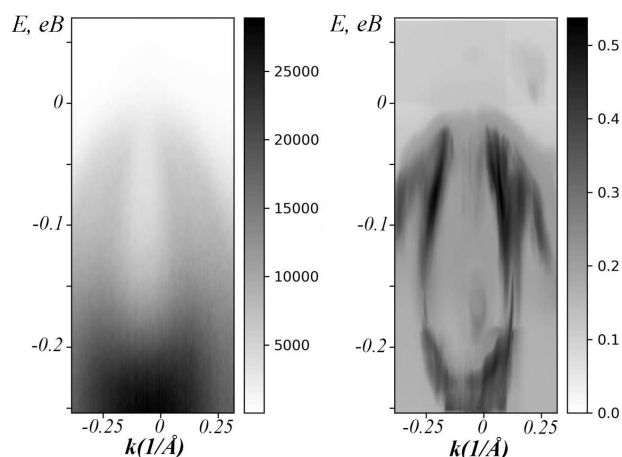


Рис. 4. Спектр ARPES у точці Z зони Брілюена в $\text{Fe}(\text{Se}, \text{Te})$, отриманий з використанням лінійної горизонтальної поляризації для температури 220 К (ліворуч) та результати її обробки за допомогою нашого методу U-Net (праворуч)

зон. У таких випадках згладжування часто призводить до втрати інформації, тому досить важко знайти компроміс між зменшенням шуму та втратою інформації щодо зон. Зокрема, за високих температур сигнал стає слабшим порівняно з шумом, що ускладнює його виділення зі спектрів або інтерпретацію потрібної інформації. Наприклад, для спектрів, отриманих при 160 К, метод кривини виявляє лише одну зону, яку можна виділити (рис. 3, с), тоді як інша зона ледве помітна. При 220 К метод кривини вже не дозволяє отримати значущу інформацію про положення зон.

На відміну від методу кривини, запропонована нами нейронна мережа ефективно візуалізує зони навіть у складних сценаріях. Як видно на рис. 3 та 4, нейронна мережа дозволяє візуалізувати зони. Високий рівень шуму значно перешкоджає здатності моделі точно класифікувати пікселі. Це видно у зниженні точності прогнозування, де ймовірність того, що піксель належить до електронної дисперсії, ледве перевищує 0,5. На противагу цьому, при нижчих температурах зазвичай отримуються ймовірності ближчі до 0,8, що вказує на набагато чіткіше розмежування між областями з дисперсією та без дисперсії.

4. Висновки

У статті пропонується розглядати проблему визначення електронної дисперсії з ARPES спектрів як

задачу семантичної сегментації. Ми провели тренування U-Net з механізмом уваги для вирішення цієї проблеми, використовуючи невеликий набір даних згенерованих спектрів та відповідні карти сегментації (2000 зразків). Обробка експериментально отриманих спектрів $\text{Fe}(\text{Se}, \text{Te})$ (отримані, використовуючи різну поляризацію випромінювання, та у температурному інтервалі 20–200 К) продемонструвала ефективність запропонованої мережі. Порівняно з методом кривини, запропонована мережа є більш стійкою до шуму і таким чином досягає кращих результатів у складних ситуаціях (кілька близько розташованих одна до одної зон, високий рівень шуму). Крім того, оскільки попередня обробка спектрів не потрібна, метод на основі U-Net може бути використаний для автоматизованої обробки великих обсягів даних.

1. M.J. Smith, J.E. Geach. *Astronomia ex machina: a history, primer and outlook on neural networks in astronomy* *Roy. Soc. Open Sci.* **10**, 221454 (2023).
2. C. Adam-Bourdarios, G. Cowan, C. Germain-Renaud, I. Guyon, B. Kegl, D. Rousseau. The Higgs machine learning challenge *J. Phys. Conf. Ser.* **664**, 072015 (2015).
3. A. Damascelli, Z. Hussain, Z.-X. Shen. Angle-resolved photoemission studies of the cuprate superconductors. *Rev. Mod. Phys.* **75**, 473 (2003).
4. J.A. Sobota, Yu He, Z.-X. Shen. Angle-resolved photoemission studies of quantum materials. *Rev. Mod. Phys.* **93**, 025006 (2021).
5. T. Valla, A.V. Fedorov, P.D. Johnson, B.O. Wells, S.L. Hulbert, Q. Li, G.D. Gu, N. Koshizuka. Evidence for quantum critical behavior in the optimally doped cuprate. *Science* **285**, 2110 (1999).
6. Z.-X. Shen, J.R. Schrieffer. Momentum, temperature, and doping dependence of photoemission lineshape and implications for the nature of the pairing potential in high- T_c superconducting materials. *Phys. Rev. Lett.* **78**, 1771 (1997).
7. R.C. Gonzalez, R.E. Woods. *Digital Image Processing. 3rd edition* (Prentice Hall, 2008).
8. P. Zhang, P. Richard, T. Qian, Y.-M. Xu, X. Dai, H. Ding. A precise method for visualizing dispersive features in image plots. *Rev. Sci. Instrum.* **82**, 043712 (2011).
9. Yu He, Y. Wang, Z.-X. Shen. Visualizing dispersive features in 2D image via minimum gradient method. *Rev. Sci. Instrum.* **88**, 073903 (2017).
10. Yo. Kim, D. Oh, S. Huh, D. Song, S. Jeong, Ju. Kwon, M. Kim, D. Kim, H. Ryu, J. Jung, W. Kyung, B. Sohn, S. Lee, J. Hyun, Ye. Lee *et al.* Deep learning-based statistical noise reduction for multidimensional spectral data. *Rev. Sci. Instrum.* **92**, 073901 (2021).

11. D. Huang, J. Liu, T. Qian, Y.F. Yang. Spectroscopic data de-noising via training-set-free deep learning method. *Sci. China Phys. Mech. Astron.* **66**, 267011 (2023).
12. J. Liu, D. Huang, Y.F. Yang, T. Qian. Removing grid structure in angle-resolved photoemission spectra via deep learning method *Phys. Rev. B* **107**, 165106 (2023).
13. H. Peng, X. Gao, Yu He, Y. Li, Y. Ji, Ch. Liu, S.A. Ekanahana, D. Pei, Z. Liu, Z. Shen, Yu. Chen. Super resolution convolutional neural network for feature extraction in spectroscopic data. *Rev. Sci. Instrum.* **91**, 033905 (2020).
14. F. Restrepo, Ju. Zhao, U. Chatterjee. Denoising and feature extraction in photoemission spectra with variational auto-encoder neural networks. *Rev. Sci. Instrum.* **93**, 065106 (2022).
15. O. Ronneberger, P. Fischer, T. Brox. U-Net: Convolutional networks for biomedical image segmentation. arXiv:1505.04597[cs.CV] (2015).
16. O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, B. Glocker, D. Rueckert. Attention U-Net: Learning where to look for the pancreas. arXiv:1804.03999v3[cs.CV] (2018).
17. T. Lin, P. Goyal, R. Girshick, K. He, P. Dollar. Focal loss for dense object detection. In: *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)* (2017), p. 2999.
18. Yu.V. Pustovit, A. A. Kordyuk. Temperature induced shift of electronic band structure in Fe(Se, Te). *Low Temp. Phys.* **45**, 1172 (2019).

Одержано 05.06.24.

Переклад на українську мову О. Войтенка

Yu. V. Pustovit,
Ye. P. Lytveniuk, Ye. D. Limarev

U-NET BASED METHOD FOR ARPES SPECTRA PROCESSING

Angle-resolved photoemission spectroscopy (ARPES) is a powerful tool for investigating the electronic structure of materials. However, resolving the electronic dispersion from ARPES spectra can be challenging due to the broadening effects, presence of different types of noise, resolution limitations, etc. This paper proposes a new approach for determining dispersion from ARPES spectra based on the U-Net neural network. The energy band extraction problem is regarded as the semantic segmentation task. We will show that the U-Net trained only with generated data can determine band structure from the experimentally obtained spectra, without prior denoising.

Keywords: ARPES, neural networks, electronic structure, U-Net.