

УДК 004.7:519.87(043.3):621.391

DOI: 10.31673/2412-9070.2020.043237

Г. Я. КІС, аспірант,
В. М. ЧЕРЕВИК, канд. техн. наук, доцент;
Державний університет телекомунікацій, Київ

ЗАСТОСУВАННЯ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ ДО СТИСНЕННЯ ДАНИХ У ПРОТОКОЛАХ ПЕРЕДАВАННЯ ЗОБРАЖЕНЬ ТА ВІДЕО

Розглянуто сучасний стан протоколів передавання даних та методів стиснення зображень та відео за допомогою штучних нейронних мереж, а саме згорткових багатошарових мереж та глибинного навчання. З огляду на останні публікації проведено порівняльний аналіз продуктивності класичних методів стиснення і методів на основі нейронних мереж, а також зроблено висновки щодо застосування цих алгоритмів у протоколі H.265 (HEVC) та можливості створення нового протоколу, повністю заснованого на нейронній мережі. Досліджено сучасні алгоритми DeepZip та DeepCoder. Надано оцінку умов переходу на нейронні мережі в перспективі.

Ключові слова: штучна нейронна мережа; стиснення даних; протокол передавання даних

ВСТУП

Стиснення зображень та відео відіграє значну роль у телекомунікації. Обмеження наявних каналів зв'язку зумовлює пошук більш ефективних підходів стиснення. Мета стиснення — найбільш компактно подати візуальний сигнал з його певною заданою якістю. Так, останній стандарт передавання відео H.265 або HEVC (*High Efficiency Video Coding*) — високоефективне кодування відео) приділяє багато уваги ефективному стисненню [1] та істотно перевищує інші стандарти за ступенем стиснення [2]. Але цей стандарт цілком засновано на класичному підході.

Водночас, останніми роками нейронні мережі та глибинне навчання, як один із найефективніших засобів тренування цих мереж, стали найкращими в численних застосуваннях оброблення зображень, зокрема розпізнаванні, семантичній сегментації, деталізації та текстуруванні. Вони довели істотний успіх у вирішенні задач оброблення зображень та комп'ютерного бачення, тому цілком зрозумілі спроби застосування сучасних архітектур нейронних мереж до задачі стиснення та його теоретичного обґрунтування.

Штучні нейронні мережі — це парадигма оброблення інформації, яка ґрунтується на подібності до біологічних нервових систем. Нині нейронні мережі ще не посіли належного місця у сфері передавання даних, але чимала низка наукових праць засвідчує інтерес науковців до цієї проблеми та висвітлює перспективи розвитку. Постає питання, який потенціал мають нейронні мережі до стиснення та коли очікувати включення нейронних мереж до телекомунікаційних стандартів. Із самого початку досліджень нейронних мереж виникали спроби їх застосування в різних галузях та, зокрема, до компресії зображень. Перші роботи стосовно застосування нейронних мереж до стиснення зображень з'явилися ще наприкінці 80-х початку 90-х років ХХ сторіччя, наприклад [4–6].

Метою цієї статті є розгляд сучасних підходів та перспектив застосування нейронних мереж до стиснення зображень та відео в наявних стандартах зв'язку.

ОСНОВНА ЧАСТИНА

Задача стиснення відео та зображень

Щоб зрозуміти корисні властивості нейронних мереж щодо стиснення зображень, розглянемо проблему стиснення взагалі.

Два базових поняття стиснення — це ступінь стиснення та якість відтворення. Необхідно якнайбільше стиснути зображення із заданим рівнем якості. Ступінь стиснення для файлів або блоків вимірюється за таким відношенням:

$$CR = \text{original image} / \text{compressed image size}.$$

Та більш зручне для потоків даних є середнє число бітів (bpp), потрібних для кодування інформації про пікселі (наприклад, такої як його яскравість та колір). Для нестиснених зображень bpp загалом характеризує кольорову модель та ступінь квантизації кольорового простору.

Якість відтворення характеризується кількома метриками порівняння зображень, зокрема піковим відношенням сигналу до шуму (peak signal-to-noise ratio) між максимумом можливого значення сигналу та потужністю шуму, що спотворює значення

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - K(i, j)|^2$$

сигналу:

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right).$$

Зрозуміло, що якісний алгоритм стиснення має зменшувати спотворення.

Як відомо, алгоритми стиснення мають дві частини — моделювання сигналу та подальше його кодування [3]. Сигнал моделюється з метою його декореляції. Наприклад, якщо дані мають лан-

цюжки однакових значень поспіль, їх можна замінити значенням, яке містить саме повторюване значення і кількість його повторів — алгоритм run length [3]. Зрозуміло, що таке перетворення дає менш скорельоване подання даних. Так, один із найбільш ефективних алгоритмів для стиснення тексту є передбачення за частковим збігом [4] (*Prediction by Partial Matching* — PPM), що обчислює наступний символ за попереднім контекстом.

Після декореляції черговим кроком є власне ентропійне кодування, наприклад арифметичне [5], яке є теоретично оптимальним у сенсі мінімуму ентропії [3].

Для стиснення зображень поняття декореляції даних теж є базовим. Якщо дані не є корельовані, можливе ефективне застосування ентропійного кодування незалежних компонентів. Загальний підхід для декореляції — це розкладання сигналу в якомусь просторовому ортогональному базисі [3]. Цей базис має бути таким, щоб мінімальна кількість компонентів несли максимальну енергію сигналу.

Теоретичною основою декореляції є теорема Карунена-Лоєва (емпіричне розкладання на власні функції). Якщо ми маємо багатовимірний дискретний сигнал, то власні вектори кореляційної матриці дають ортонормований базис із двома важливими властивостями. Сигнал у цьому базисі буде декорельований (за означенням кореляція між ортогональними компонентами відсутня) та похибка відтворення сигналу є мінімальною у сенсі відкидання компонент із малою енергією. Більш того, використовуючи цю теорему, можна задавати ступінь стиснення або потрібну якість відтворення в алгоритмі.

З практичного погляду пряме застосування перетворення Карунена-Лоєва не є обчислювально-ефективним, але можна скористатись його наближеннями.

Методи стиснення на основі декореляційних перетворень називаються трансформційним кодуванням [19]. Найбільш ефективними вважаються дискретне косинусне перетворення (стандарт JPEG) [6], упроваджене 1972 року Насіром Ахметом [7], та вейвлет-перетворення (стандарт JPEG-2000) [8]. Коефіцієнти перетворень мають цілком зрозумілий фізичний зміст просторових частот та

можуть піддаватися подальшій квантизації для більш оптимального подання компонентів, які є менш важливі для сприйняття людиною (рис. 1).

Також для передбачення може застосовуватися інтерполяція сигналу, наприклад диференційно-імпульсна кодова модуляція [9] між сусідніми блоками зображення. Зрозуміло, що сусідні блоки на зображенні є подібними, але більш раціонально кодувати тільки різницю між сусідніми блоками кодування. Стандарт JPEG [6] визначає сім типів диференційного передбачення даних блока за оточуючими трьома блоками. Взагалі, предиктивні методи більш застосовуються до стиснення без втрат.

Порівняльні результати стиснення [9] з фіксованим рівнем PSNR та MOS (середнє суб'єктивне оцінювання) наведено в табл. 1.

Таблиця 1
Погіршення стиснення, %, JPEG/JPEG-2000
відносно рівня HEVC [9]

Metric	JPEG	JPEG-2000 4:2:0	JPEG-2000 4:4:4
PSNR	-61,63	-20,26	-46,88
MOS	-43,10	-30,96	-15,98

Останній Стандарт HEVC дає ефективнішу схему компресії зображень, яка ще більш використовує подібність сусідніх блоків та застосовує інтерполяцію (intra coding). Так, у 2014 році Белардом на основі стандарту HEVC було розроблено формат BPG (better portable graphics) [12] для застосування в Інтернеті як альтернатива JPEG, який є набагато ефективніший за інші стандарти.

Застосування штучних нейронних мереж до стиснення

Невеликий огляд сучасного стану відкритих стандартів, який було надано у попередньому розділі, показує, які властивості нейронних мереж можливо застосувати до стиснення.

По-перше, мінімізація спотворень алгоритмом стиснення є дуже подібною до методів глибинного навчання, де мінімізується цільова функція (loss function). Та, по-друге, нейронні мережі ефективно узагальнюють дані, відокремлюючи у такий спосіб незалежні компоненти вхідного сигналу. Цією властивістю можна скористатися для декореляції даних [11].

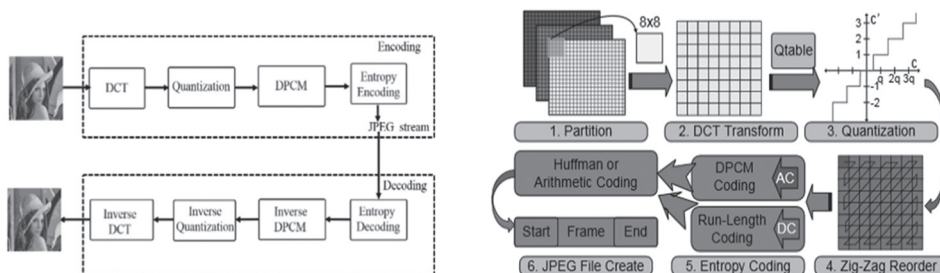


Рис. 1. Загальна схема алгоритму JPEG та більш детальна схема стиснення [12 DeepZip]

Виокремлюють два підходи для стиснення за допомогою нейронних мереж: у першому нейронна мережа використовується як частина наявного алгоритму (гібридне кодування), а у другому підході нейронна мережа дає стисле подання даних (стискальна мережа).

У першому підході нейронні мережі розглядалися як більш ефективні методи для моделювання даних з погляду їхньої адаптивності та можуть бути використані як частини стандартних алгоритмів.

Наведемо кілька прикладів, характерних для першого підходу.

Нейронна мережа може сегментувати зображення на семантично важливі регіони (обличчя людини тощо) та задавати різний ступінь стиснення для різних регіонів. Отже, неважливі для людського сприйняття частини зображення можуть бути стиснені більш істотно.

Також нейронні мережі можуть використовуватися для більш ефективного передбачення наступного елемента в потоці даних за контекстом, наприклад як у RPM. Так, рекурентні нейронні мережі дуже добре моделюють контекст та часові послідовності. Вони успішно застосовуються для розпізнавання мови та рукописного тексту. Нещодавно було запропоновано використати такі мережі для стиснення даних, як DeepZip [12]. Автори доводять, що застосування категорійної крос-ентропії (categorical cross-entropy) є оптимальною функцією втрат до задачі стиснення без втрат (lossless compression). У роботі розглянуто три архітектури мережі: щільно зв'язана (fully connected, FC), довга короткочасна пам'ять (long-short-term memory, LSTM) та вентильні рекурентні вузли (Gated recurrent units, GRU) з подальшим порівнянням результатів із класичними алгоритмами. Результати досліджень наведено в табл. 2. Бачимо, що нейронні мережі дають не гірший результат за класичні підходи.

Таблиця 2

Результати компресії, МВ, для кількох відомих колекцій зображень та алгоритмів [12]

Dataset	Seq. Length	Gzip	BSC	DeepZip		
				FC	biGRU	LSTM-multi
H.chr1	249M	60,58	50,43	49,37	48,80	48,56
C. E. chr 1	15M	4,03	3,49	3,81	3,58	4,02
C. E. genome	100M	26,97	23,38	23,41	23,13	23,41
text8	100M	33,05	20,95	25,49	23,37	26,71
PhiX Quality	100M	6,22	4,38	4,58	4,35	4,79

Для зображень, як і в разі кодування текстів, нейронні мережі можуть бути використані для

передбачення значень коефіцієнтів перетворень, узагальнюючи диференційно-імпульсну модуляцію та інтра-кодування.

З появою багатшарових мереж, які мають кілька прихованих шарів, теоретично було доведено, що багатшарова мережа з більш ніж одним прихованим шаром може наблизити будь-яку неперервну функцію із заданою точністю [15]. Ця властивість дає можливість моделювати довільний алгоритм у термінах нейронної мережі. Так було запропоновано схему, яка повторює роботу алгоритму стиснення. Можна стверджувати [23], що вперше в 1988 році було сформульовано традиційні кроки зі стиснення [16], зокрема унітарне просторове перетворення зображення, квантизацію та кодування квантизованих коефіцієнтів як задачу оптимізації певної цільової функції. Дійсно, якщо застосувати належну функцію втрат, можливо побудувати алгоритм стиснення. Такі стискальні мережі було запропоновано вже 1989 року [17] — трансформаційна мережа з трьома шарами та навчанням за допомогою зворотного поширення похибки, яка мала б давати коефіцієнти ДКП.

Розглянемо властивості згорткових нейронних мереж щодо декореляції даних. Згорткова мережа, запропонована Яном Лекуном у 1988 році, успішно застосовується до аналізу візуальних зображень. Проаналізуємо її роботу. Згорткові шари комбінуються із шарами підвибірки (пулінг), які вибирають найбільш важливі компоненти, ущільнюючи мережу шар за шаром. Після кількох проходжень згортки та ущільнення дані перебудовуються від пікселів із високим розділенням до більш абстрактних ознак. Як правило, на кожному наступному шарі збільшується кількість каналів і зменшується розмір зображення в кожному каналі. Урешті-решт залишається набір каналів, що зберігають невелику кількість даних, які інтерпретуються як найабстрактніші поняття, виявлені з початкового зображення. Цю мережу можна інтерпретувати як екстрактор особливостей, що перетворює зображення в простір із компактним поданням, що вельми корисно для стиснення [13].

Бачимо, що вибір компонентів нейронною мережею дуже нагадує перетворення Карунена-Лоєва.

Більш доцільною виявилась схема, де для кодування зображень застосовують більш вузькі шари прихованих нейронів, які прагнуть знайти особливості вхідного зображення, а зворотна мережа реконструює зображення на кінцевому шарі (рис. 2). Така архітектура в нейронних мережах має назву автоенкодера. Останніми роками з'явилася низка робіт, в яких вивчався цей підхід та було досягнуто істотних результатів. Як приховані шари було запропоновано рекурентні [19] та більш прості згорткові шари [20].

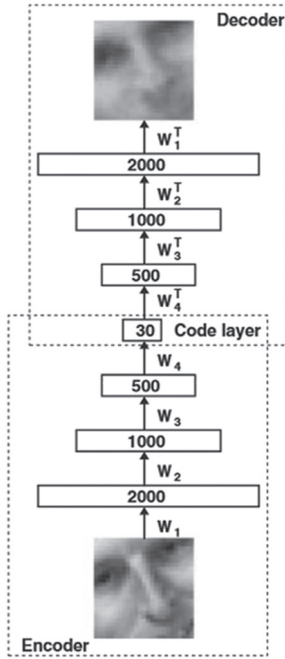


Рис. 2. Схема автоенкодера до задачі стиснення даних [15]

Останній підхід виявився дуже перспективним, цей автоенкодер подібний до JPEG-2000 за якістю та ступенем стиснення. Він може обробляти зображення високої роздільної здатності. Результати стиснення автоенкодерів надано в роботі [20] та зображено на рис. 3, де бачимо, що ці мережні алгоритми перевищують якість класичних підходів (результат Тодерічі подано без фінального ентропійного кодування).

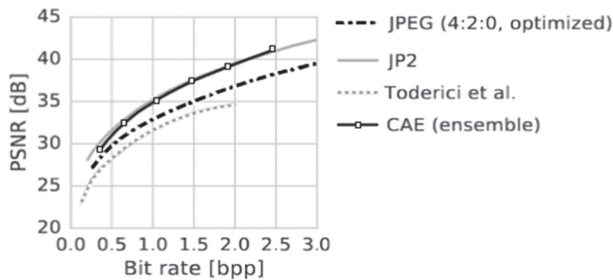


Рис. 3. Порівняння алгоритмів JPEG, JPEG-2000, Тодерічі та Теїс за PSNR на Kodak PhotoCD наборі зображень

Нещодавно було запропоновано застосувати дворівневу систему стиснення [21] (рис. 4). Спочатку використовується автоенкодер, вихід енкодера передається в мережу передбачення за контекстом для ентропійного моделювання (за авторами hyperprior network). Та остаточно ця статистична модель застосовується в арифметичному кодері. Автори повідомляють, що цей підхід дає найкраще стиснення серед усіх інших алгоритмів, які використовують глибинне навчання, на 15,8% та краще за JPEG на 59,8%.

Стиснення відео

Більшість концепцій щодо компресії відео можливо побачити в стандарті передавання відео MPEG. Перші спроби зі створення цього стандарту розпочались у 1988 році, та зараз маємо останню версію MPEG-H (HEVC або H.265). Головною перевагою відносно інших стандартів було те, що це був відкритий стандарт.

Істотна відмінність відео від статичного зображення — це наявність часової кореляції між сусідніми кадрами у відео. Компенсація руху є одним із популярних методів зменшення цієї кореляції. Блоки зображень кодуються як зсунена копія з попереднього кадру. На основі нейронних мереж було розроблено низку підходів для обчислювання руху блоків.

Також різні частини кадру мають різну інформаційну цінність та активність. Зрозуміло, що менш активні блоки можна кодувати з більшою квантизацією. Отже, блоки кадру можуть бути подані у вигляді квадро-дерев із приблизно однаковою внутрішньоблоковою варіацією. Ця структура може зберігатися деякий проміжок часу, що також може бути застосовано для кодування. Нейронні мережі добре відповідають задачам класифікації і теж можуть бути використані для вибору належного типу кодування блока [22].

Гібридні схеми застосування нейронних мереж у підході HEVC досліджуються в різних підсистемах, зокрема аналіз руху, знаходження кореляцій

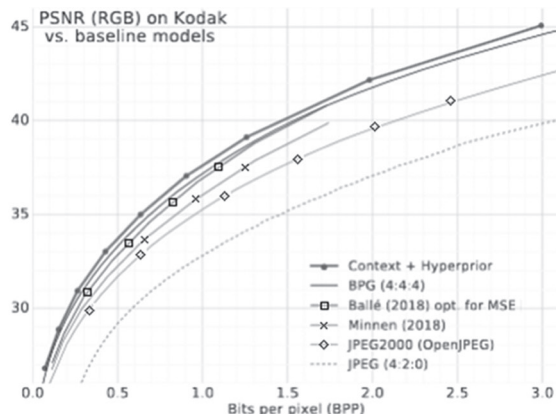
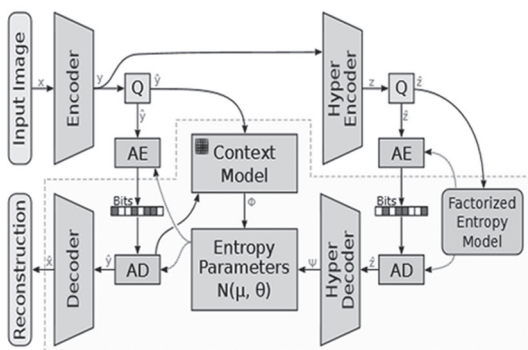


Рис. 4. Дворівнева схема автоенкодерів та порівняння продуктивності алгоритмів стиснення згідно з [21]

та інтра- і інтер-передбачення, статистичне моделювання контексту, зменшення блочного ефекту (loop filters — кільцеві фільтри). Огляд різних архітектур мереж до цих задач можна знайти в [23].

Також пропонуються схеми, які повністю засновані на нейронних мережах. Наприклад, можливо розглянути DeepCoder [24], який було запропоновано 2017 року. Тут усі блоки зображення кодуються за допомогою передбачення за мапою особливостей. Похибки інтра- та інтер-передбачення також кодуються допоміжною мережею. Після квантизації мапи передбачених значень (feature map) та нев'язки кодуються алгоритмом Хафмана.

Нещодавно (2019 рік) було запропоновано іншу мережу Deep Learning Video Coding (DLVC) [25], яка, за ствердженням авторів, стискає на 39,6% краще, ніж HEVC. Також автори відкрили свій код для інших дослідників. Автори розмірковують, чи можуть нейронні мережі замінити класичні алгоритми. На їх думку, це може статися у двох випадках — глибинне навчання безперечно стане кращим за класичні алгоритми, та якщо в HEVC усі структурні блоки будуть поступово замінені нейронними мережами, виконуючи якусь певну функцію. Також автори зазначають, що мережі мають набагато більшу обчислювальну складність за класичні підходи, і це стримує їх упровадження до стандартів.

Висновки

Алгоритми на основі нейронних мереж досягли якості класичних алгоритмів, але вони потребують набагато більше обчислювальних ресурсів.

З розвитком спеціалізованих чіпів для паралельних обчислень для вбудованих систем, алгоритми з використанням згорткових нейронних мереж можуть стати більш доцільними до використання.

Так, незважаючи на стрімке поширення стандарту JPEG, стандарт JPEG-2000 ще не став популярним за 20 років, зокрема кількох відомих застосувань, попри його явні переваги в якості стиснення. Те саме стосується і протоколу H.265 (HEVC), який є розширенням протоколу H.264 з міркувань сумісності обладнання та наявних даних.

Водночас нейронні мережі дійсно показують кращі результати, але нині їх важко стандартизувати, щоб отримувати очікуваний результат за різних архітектур мереж.

Отже, можна дійти такого висновку: якщо очікувати впровадження нейронних мереж, то воно може відбуватись як розширення наявних стандартів. Розроблення самостійного стандарту є маловірогідним у найближчий час. Мережі буде застосовано для більш ефективного моделювання

контексту, адаптивному до кодування та вибору рівня квантизації.

Що стосується стиснення відеоданих, алгоритми на основі нейронних мереж ще не є досконалими та менш ефективні, ніж HEVC. Можна очікувати подальше теоретичне вивчення та спроби побудови успішних мережних архітектур.

У майбутньому можна передбачити збільшення потреби в якості передавання відео, що стикнеться з обмеженням класичних підходів. Сьогодні ж активно йде розроблення спеціалізованих процесорів для паралельного оброблення даних та нейронних мереж. Ці два фактори вказують, що нейронні мережі мають бути впроваджені в промислових стандартах передавання даних.

Список використаної літератури

1. *Overview of the High Efficiency Video Coding (HEVC) Standard* / Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas Wiegand // *IEEE Trans. on Circuits and Systems for Video Technology*. Dec. 2012.

2. *Sullivan G. J. Comparison of the Coding Efficiency of Video Coding Standards – Including High Efficiency Video Coding (HEVC)* // *IEEE Trans. on Circuits and Systems for Video Technology*. 2012.

3. *Khalid Sayood. Introduction to Data Compression: 4th Edition*. October 2012.

4. *Cleary J. G., Witten I. H. Data compression using adaptive coding and partial string matching* // *IEEE transactions on communications*. April 1984. Vol. COM-32, No. 4. P. 396.

5. *Ian H. Witten, Radford M. Neal, John G. Cleary. Arithmetic coding for data compression* // *Commun. ACM*. June 1987. Vol. 30, No. 6. P. 520–540.

6. *Wallace G. K. The JPEG still picture compression standard* // *Comm. ACM*. April 1991. Vol. 34. P. 30–44.

7. *High-Performance Compression of Visual Information - A Tutorial Review - Part I: Still Pictures*. / O. Egger, P. Fleury, T. Ebrahimi, M. Kunt // *Proceedings of the IEEE*. June 1999. Vol. 87, No 6.

8. *Ahmed N. How I Came Up With the Discrete Cosine Transform*. *Digital Signal Processing*. 1(1): 4–5.

9. *U. S. patent 2605361*, C. Chapin Cutler, *Differential Quantization of Communication Signals*, filed June 29, 1950, issued July 29, 1952.

10. *Rabbani M., Joshi R. An overview of the JPEG2000 still image compression standard* // *Signal Processing: Image Communication* 17. Jan. 2002. P. 3–48.

11. *AhG4: Subjective evaluation of HEVC intra coding for still image compression* / P. Hanhart, M. Re-rabek, P. Korshunov, T. Ebrahimi // *JCT-VC*. 2013.

12. *Bellard F. BPG Image format, 2014* [Електронний ресурс] URL:

https://bellard.org/bpg/bpg_spec.txt

13. **Hinton G. E., Salakhutdinov R. R.** *Reducing the Dimensionality of Data with Neural Networks* // SCIENCE. 2006. Vol. 313. P. 504–507.
14. **DeepZip: Lossless Data Compression using Recurrent Neural Networks** arXiv:1811.08162v1 [cs.CL] / Mohit Goyal, Kedar Tatwawadi, Shubham Chandak, Idoia Ochoa. 20 Nov. 2018.
15. **Schalkoff R. J.** *Artificial Neural Networks*. McGraw-Hill, 1997.
16. **Chua L., Lin T.** *A neural network approach to transform image coding* // International Journal of Circuit Theory and Applications. 1988. Vol. 16, No. 3. P. 317–324.
17. **Image data compression using a neural network model** / N. Sonehara, M. Kawato, S. Miyake, K. Nakane // Proc. IJCNN. 1989. Vol. 2. P. 35–41.
18. **Cotrell G. W., Munro P., Zipser D.** *Image compression by back propagation: an example of extensional programming* // Models of cognition: rev. Of cognitive science. 1989. Vol. 1, No. 208. P. 1.
19. **Full resolution image compression with recurrent neural networks** / G. Toderici, D. Vincent, N. Johnston [et al.]. 2016. arXiv:1608.05148v1.
20. **Lossy Image Compression with Compressive Autoencoders** / L. Theis, W. Shi, A. Cunningham, F. Huszár. arXiv:1703.00395 [stat.ML]. 2017.
21. **Minnen D., Ballé J., Toderic G.** *Joint Autoregressive and Hierarchical Priors for Learned Image Compression*. arXiv:1809.02736v1 [cs.CV] 8 Sep. 2018.
22. **Vigliano Daniele, Raffaele Parisi, Aurelio Uncini.** *Video compression by Neural Network* // Studies in Fuzziness and Soft Computing. 2005.
23. **Image and Video Compression with Neural Networks** / Siwei Ma, Xinfeng Zhang, Chuanmin Jia [et al.] // A Review arXiv:1904.03567v2 [cs.CV] 10 Apr. 2019.
24. **DeepCoder: A deep neural network based video compression** / Chen, H. Liu, Q. Shen, T. Yue [et al.] // Visual Communications and Image Processing (VCIP), 2017. IEEE, P. 1–4.
25. **Deep Learning-Based Video Coding** / Dong Liu, Yue Li, Jianping Lin [et al.] // A Review and A Case Study. ArXiv:1904.12462v1 [cs.MM] 29 Apr. 2019.

В. Ч. Черевик, Г. Я. Кис

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ СЖАТИЯ ДАННЫХ В ПРОТОКОЛАХ ПЕРЕДАЧИ ВИДЕО

Рассмотрены современные подходы в протоколах передачи данных в контексте сжатия изображений и видео с помощью нейросетей, а именно сверточных многослойных сетей и методов глубинного обучения. На основе недавних публикаций проведен сравнительный анализ производительности классических методов сжатия и методов на основе нейросетей, в частности, были рассмотрены алгоритмы DeepZip и DeepCoder. Анализ позволяет сделать выводы относительно использования нейросетей в протоколе H.265 (HEVC) и возможности создания нового протокола, полностью основанного на нейросетях. Дана оценка условий перехода в перспективе на нейронные сети.

Ключевые слова: искусственная нейросеть; сжатие данных; протокол передачи видео.

V. M. Cherevyk, G. Ya. Kis

ARTIFICIAL NEURAL NETWORK APPLICATIONS FOR DATA COMPRESSION IN VIDEO DATA TRANSFER PROTOCOLS

The article describes the current state of data transfer protocols and methods of image and video compression through the use of artificial neural networks, namely convolutional multilayer networks and deep structured learning. Based on recent publications, a comparative analysis of the performance of classical compression methods and methods based on neural networks was performed.

The most effective are those compression methods which are based on decorrelation transforms, namely discrete cosine (JPEG standard) and wavelet (JPEG-2000 standard) transforms. The transform coefficients have a well-understood physical content of spatial frequencies and can be further quantized for a more optimal representation of components that are less important for human perception. The HEVC standard guarantees a more efficient image compression scheme that further takes advantage of the similarity of adjacent blocks and uses interpolation (intra coding). Based on the HEVC standard, the BPG (better portable graphics) format was developed to be used on the Internet as an alternative to JPEG, which is much more efficient than other standards.

An overview of the current state of open standards, provided in the article, gives an explanation of what properties of neural networks can be applied to image compression. There are two approaches towards the compression using neural networks: in case of the first approach neural network is used as a part of an existing algorithm (hybrid coding), and in case of the second approach the neural network gives a concise representation of the data (compression network).

The final conclusions were made as regards to the application of these algorithms in H.265 protocol (HEVC) and the possibility of creating a new protocol which is completely based on the neural network. Protocols using neural network show better results during image compression, but are currently hard to be subjected to standardization in order to obtain the expected result in case of different network architects.

We may expect and predict an increase in the need for video transmission in the future, which will bump into the imitating nature of classical approaches. At the same time, the development of specialized processors for parallel data processing and implementation of neural networks is currently underway. These two factors indicate that neural networks must be embedded into the industrial data standards.

Keywords: telecommunication data transfer protocol; artificial neural networks; video and image compression.