Uniwersytet Morski w Gdyni, Wydział Elektryczny ORCID: 1. 0000-0003-3517-4046; 2.0000-0003-2941-2415

Rozpoznawanie i rekonstrukcja obrazów jako problem odwrotny z zastosowaniem systemu uczenia maszynowego

Streszczenie. Rekonstrukcja i rozpoznawanie obrazów jest typowym problemem występującym w wielu systemach przetwarzania obrazów. Zagadnienie to można sformułować jako rozwiązanie problemu odwrotnego. W artykule zaprezentowano autorski model systemu uczenia maszynowego, który może być wykorzystany do rekonstrukcji i rozpoznawania obrazów na podstawie ich liniowych projekcji.

Abstract. Image reconstruction and recognition is a common problem found in many image processing systems. This problem can be formulated as a solution to the inverse problem. The article presents the author's model of a machine learning system that can be used for the reconstruction and recognition of images based on their linear projections. (Image Recognition and Reconstruction as Inverse Problem, Using Machine Learning System).

Słowa kluczowe: systemy uczenia maszynowego, rekonstrukcja i rozpoznawanie obrazów, problem odwrotny **Keywords**: machine learning systems, image reconstruction and recognition, inverse problem.

Wstęp

Systemy rekonstrukcji i rozpoznawania obrazów wykorzystujące klasyczne metody są obecnie uzupełniane i częściowo zastępowane systemami sztucznej inteligencji do których zaliczane są także algorytmy uczenia maszynowego. Przegląd metod uczenia maszynowego w rozpoznawaniu obrazów ze szczególnym uwzględnieniem zastosowań medycznych można znaleźć w pracach [1-4]. Warto zauważyć, że obecne systemy rekonstrukcji i rozpoznawania obrazów oparte są najczęściej na algorytmach optymalizacji z ograniczeniami przy zastosowaniu odpowiednio dobranych metod regularyzacji [5].

Ostatnio szeroko stosowane algorytmy głębokiego zainteresowania uczenia spowodowały renesans sztucznymi sieciami neuronowymi oraz różnego rodzaju ich aplikacjami. Większość znanych algorytmów głębokiego uczenia posiada swoje implementacje w postaci sztucznych sieci neuronowych (Artificial Neural Networks) oraz sztucznych sieci neuronowych głębokiego uczenia (Deep Learning Neural Networks) wyuczonych na podstawie zbioru danych uczących poprzez minimalizacje funkcji strat. W tym kontekście podejście do głębokiego uczenia może być postrzegane jako szczególny przypadek z zakresu teorii metod optymalizacji. Standardowe typy sieci neuronowych głębokiego uczenia obejmują wielowarstwowe perceptrony (MLP), splotowe sieci neuronowe (CNN), rekurencyjne sieci neuronowe (RNN) i generatywne sieci kontradyktoryjne (GAN) [6-9]. Analizując dostępne publikacje można stwierdzić, że optymalna topologia sieci i technologia jej implementacji nie została jeszcze określona. Nie jest także dobrze zbadany związek topologii sieci z jej wydajnością [10]. Niemniej, można stwierdzić, że SNN stanowią uniwersalne modele algorytmiczne i fizyczne stosowane w systemach inteligencji obliczeniowej. Jednym z podtypów SNN są sieci neuronowe typu Hopfielda, które są zarówno modelami fizycznymi jak i algorytmami stosowanymi w obliczeniach neuronowych. We wcześniejszych pracach zaproponowaliśmy rozszerzony model sieci neuronowej typu Hopfielda zdefiniowany przez następujące równanie [11]:

(1)
$$\dot{x} = (\eta W - w_0 \mathbf{1} + \varepsilon W_s) \boldsymbol{\theta}(x) + I_d$$

gdzie:*W*— skośnie symetryczna macierz ortogonalna; *W*_s— macierz rzeczywista symetryczna; **1**—macierz jednostkowa; $\theta(x)$ —wektor funkcji aktywacji; *I*_d—wektor wejściowy; ε , *w*₀, η —parametry

Równanie (1) w stanie równowagi sieci przyjmuje postać:

(2)
$$(\eta W - w_0 \mathbf{1} + \varepsilon W_s) \boldsymbol{\theta}(\mathbf{x}) + I_d = \mathbf{0}$$

Równanie (2) stanowi bazę dla uniwersalnych modeli uczenia maszynowego opartych na przekształceniach biortogonalnych, umożliwiających realizację typowych funkcji systemów uczenia się. Jedną z tych funkcji jest implementacja pamięci asocjacyjnych. Zastosowanie systemu do rekonstrukcji i rozpoznawania zniekształconych/zaszumionych obrazów z wykorzystaniem pamięci asocjacyjnej było opisane szerzej w pracach [11-13].

W bieżącej pracy przebadana została implementacja systemu uczenia maszynowego do rozwiązywania problemów odwrotnych (Inverse Problem). W ogólnym przypadku zagadnienie odwrotne polega na wyznaczeniu przyczyny (danych wejściowych) na podstawie skutku (danych wyjściowych). W przedstawionych w pracy doświadczeniach oryginalny obraz był przetwarzany przez liniowy operator macierzowy, którego wymiar nie był kwadratowy. Nie istniał zatem operator odwrotny w sensie algebry macierzowej. Odpowiednio zaprojektowany system uczenia maszynowego wykonywał rekonstrukcję oryginalnego obrazu na podstawie jego projekcji.

System uczenia maszynowego do przetwarzania obrazów

Rozważamy zbiór *N* czarno-białych obrazów reprezentowanych przez *k* wierszy i *l* kolumn, czyli zbiór $(k \cdot l)$ pikseli o różnych odcieniach szarości. W przypadku analizy wektorowej każdy obraz jest przekształcany przez konkatenacje w wektor kolumnowy x_i $(k \cdot l \times 1)$, i = 1, ..., N. Tak więc zbiór *N* obrazów jest reprezentowany przez następującą macierz:

(3) $X = [x_1, x_2, ..., x_N], \dim x_i = k \cdot l = 2^q, q = 3, 4, ...,$

Zbiór zniekształconych obrazów reprezentowany jest przez macierz:

(4)
$$X^{(s)} = \left[x_1^{(s)}, x_2^{(s)}, \dots, x_N^{(s)} \right]$$

Łatwo zauważyć, że zbiór treningowy można zapisać następująco:

(5)
$$S = \{x_i, x_i^{(s)}\}_{i=1}^{N}$$

Zbiór *S* tworzy odwzorowanie $F(\cdot)$ zdefiniowane przez następujące właściwości:

$$(6) x_i = F(x_i)$$

oraz

(7)
$$x_i^{(s)} \stackrel{F}{\to} x_i, i = 1, 2, ..., N.$$

W ten sposób odwzorowanie *F* jest zaimplementowane jako system uczenia maszynowego do rekonstrukcji obrazu. Strukturę realizującą odwzorowanie $F(\cdot)$ określoną równaniami (6) i (7) można otrzymać jako rozwiązania równania równowagi (2). Zatem dla parametrów $w_0 = 2$, $\varepsilon = 1$ w równaniu 2 otrzymujemy:

(8)
$$(W_{2^k} - 2 \cdot 1 + W_s)m_i + x_i^{(s)} = 0$$

gdzie: $W_{2^k}^2 = -1$, skośnie-symetryczna macierz ortogonalna

Uzyskujemy zatem N rozwiązań:

(9)
$$\boldsymbol{m}_{i} = (2 \cdot 1 - \boldsymbol{W}_{s} - \boldsymbol{W}_{2^{k}})^{-1} \boldsymbol{x}_{i}^{(s)}, i = 1, ..., N$$

gdzie

$$(10) W_s = M(M^T M)^{-1} M^T$$

(11)
$$M = \{m_1, m_2, ..., m_N\}$$

jest macierzą widmową wektorów x_i

(12)
$$m_i = \frac{1}{2} (W_{2^k} + 1) x_i$$

tak więc

(13)
$$x_i = (-W_{2^k} + 1)m_i, i = 1, ..., N$$

Wektory x_i stanowią tzw. wektory systemowe $u_i \equiv x_i, i = 1, ..., N$.

Równanie (9) definiuje transformację biortogonalną $T_s(\cdot)$:

(14)
$$\boldsymbol{m}_i = T_s(\boldsymbol{x}_i^{(s)})$$

a równanie (12) transformację ortogonalną:

$$(15) x_i = T^{-1}(\boldsymbol{m}_i)$$

Transformacje $T_s(\cdot)$ oraz $T^{-1}(\cdot)$, realizujące odwzorowanie $F(\cdot)$ mają strukturę blokową przedstawioną na rysunku 1[11]. Transformacja ortogonalna $T_s(\cdot)$, wykorzystująca rodzinę macierzy Hurwitza–Radona, pozwala na wyznaczenie widm Haara–Fouriera wektorów x_i .



Rys. 1. Struktura system uczenia maszynowego do przetwarzania obrazów

Struktura z rysunku 1 służy jako estymator widma $\{ \hat{m}_i \}$:

(16)
$$\widehat{\boldsymbol{m}}_i = T_s\left(\boldsymbol{x}_i^{(s)}\right), i = 1, \dots, N$$

W systemie, ze względu na iteracyjny charakter pętli sprzężenia zwrotnego, uzyskuje się zbieżność wektorów:

(17)
$$\widehat{\boldsymbol{m}}_i \to \boldsymbol{m}_i$$

(18)
$$\widehat{y}_i \to x_i, i = 1, \dots, N.$$

Zbieżność procesu uzyskuje się po *L* iteracjach przy czym liczba *L* jest różna dla różnych procesów

rekonstrukcji. Ponadto należy zauważyć, że dla obrazu wejściowego $z \neq x_i$, i = 1, ..., N, wyjście systemu jest dane przez superpozycję wektorów wejściowych:

(19)
$$F(\mathbf{z}) = \sum_{i=1}^{N} \alpha_i \mathbf{x}_i, \ \alpha_i \in \mathbb{R}$$

Wektory systemowe $u_i = x_i$ tworzą centra atrakcji a system z rysunku 1 realizuje funkcję rozłożonej pamięci asocjacyjnej.

Rozpoznawanie i rekonstrukcja obrazów jako problem odwrotny

Przedstawione w poprzednim rozdziale modele rekonstrukcji obrazu bazują na dostępności zbiorów uczących *S*, zawierających oryginalne i uszkodzone wzorce. Alternatywnie, ogólny model rekonstrukcji obrazu może być zapisany równaniem:

$$(20) Ax = \widetilde{y}$$

gdzie: A— znany operator liniowy np. operator macierzowy, x—oryginalny obraz; \tilde{y} —obserwowany zdegenerowany obraz

Zgodnie z równaniem (20) rekonstrukcja obrazu prowadzi do rozwiązania problemu odwrotnego. Większość znanych z literatury rozwiązań równania (20) wykorzystuje metody optymalizacyjne [5,14], na przykład:

(21)
$$\min_{x} \|\widetilde{y} - Ax\|_{2}^{2}, \text{ s.t. } x \in K$$
$$\min_{x} \|\widetilde{y} - Ax\|_{2}^{2} + \beta R(x)$$

K—zbiór dopuszczalnych rozwiązań; R(x)—regularyzator ; β —parameter regularyzacji

Jak wspomniano powyżej, różne typy sieci neuronowych są obecnie wykorzystywane do rozwiązywania odwrotnych problemów w obrazowaniu, w tym rekonstrukcji obrazu. Wiele podejść do tego problemu można znaleźć w pracach przeglądowych [14-16]. Wykorzystanie modelu uczenia maszynowego pokazanego na rysunku 1 do rozwiązania równania (20) prowadzi do rozwiązania następującego problemu:

$$F(\mathbf{x}): \mathbf{A}\mathbf{x} = \hat{\mathbf{y}}$$

 $x = F^{-1}(\widetilde{y})$

gdzie: $A - (m \times n)$ znana macierz rzeczywista, $m \neq n$; $\tilde{y} - (m \times 1)$ macierz rzeczywista; $x - (n \times 1)$ macierz rzeczywista; $m + n = 2^q$, q = 3, 4, ...

Przypadek m = n nie jest rozważany. Generacja zbioru uczącego $S = \{x_i, y_i\}_{i=1}^N$ dla równania (20) jest dana wzorem:

(23)
$$Ax_i = y_i, i = 1, 2, ..., N$$

gdzie x_i , i = 1, 2, ..., N są postaciami wektorowymi oryginalnych obrazów treningowych. Zakładając, że macierz projekcji $A(m \times n)$ w równaniu (22) jest macierzą losową, obrazy y_i zbioru uczącego stają się losowymi wektorami. Przykładowy obraz oraz jego projekcję przedstawiono na rysunku 2.

Wektorowa transformacja wybranego obrazu przyjmuje postać:

$$(24) Ax = y$$

(22)

gdzie: dim $A = (m \times n), m > n$.

Tworząc wektory systemowe u_i o postaci:

(25)
$$\boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{y}_i \\ \boldsymbol{x}_i \end{bmatrix}, i = 1, \dots, N$$

strukturę odwzorowania odwrotnego (22):

(26)
$$x_i = F^{-1}(y_i), i = 1, ..., N$$

podano na rysunku 3 a, b.



Rys. 2. Oryginalny obraz oraz jego projekcja (transformacja)



 \hat{x}_i —estymator obrazu x_i ,

 \widetilde{y}_i —projekcja zdegenerowanego obrazu x_i .

Rys. 3. Struktura systemu realizującego transformację odwrotną (a) y_i – projekcja niezdegenerowanego obrazu

(b) $\widetilde{\mathbf{y}}$ – projekcja zdegenerowanego obrazu

Należy zauważyć, że transformacja biortogonalna $T_s(\cdot)$ i transformacja ortogonalna $T^{-1}(\cdot)$ na rysunku 3 są określone odpowiednio równaniami (14) i (15):

(27)
$$\boldsymbol{m}_i = T_s \left(\begin{bmatrix} \boldsymbol{y}_i \\ \boldsymbol{0} \end{bmatrix} \right)$$

(28)
$$u_i = T^{-1}(m_i)$$

gdzie: u_i —wektory systemowe (25).

W systemie przedstawionym rysunku na zniekształcone rzuty obrazów $\widetilde{y}_i, i = 1, \dots, N$ ulegają rekonstrukcji. Aby zilustrować właściwości systemu rekonstrukcji, przedstawionego na rysunku 3. wygenerowano korzystając z równania (24) zbiór uczący S. Do generacji zbioru użyto dziewięciu obrazów x_i , i =1,...,9. Przykładowe przekształcenie obrazu pokazano na rysunku 4. W układzie pokazanym na rysunku 3b uzyskujemy:

(29)
$$||x - \tilde{x}||_2^2 = 0$$

Podsumowując, rysunki 1 i 3 przedstawiają systemy rekonstrukcji obrazu, które wykorzystują pamięć asocjacyjną do rozpoznawania uszkodzonych wzorców. Warto jednak zauważyć, że system na rysunku 3 realizuje transformację odwrotną i rozwiązuje zadania optymalizacyjne ograniczone obrazami przechowywanymi w rozłożonej pamięci. Zbieżność ciągu rekurencyjnego zachodzi dla $N < \frac{1}{2}(n+m)$.



Rys. 4. Przykładowa rekonstrukcja (F (·) – system z rysunku 3b)

Ponadto system ten umożliwia również rozwiązywanie równań liniowych (24) przy użyciu losowej postaci wektorów uczących x_i w równaniu (23). Nową strukturę systemu nauczania maszynowego dla rozwiązania problemu odwrotnego (20), uzyskuje się zatem generując zbiór uczący z wykorzystaniem wektorów losowych x_i w równaniu (23). Tak więc x_i , i = 1, ..., N jest losową formą wektorów (obrazów) treningowych, przy czym:

$$(30) N = m$$

oraz y_i , i = 1, ..., N są projekcjami obrazów x_i . Struktura systemu zadana jest zależnościami (8)-(16) przy czym macierz symetryczna $W_s(10)$ musi być regularyzowana przez przepis regularyzujący dla macierzy pseudoodwrotnej:

$$W_s = \boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M} + \gamma \boldsymbol{1})^{-1}\boldsymbol{M}^T$$

gdzie: $\gamma > 0$

Wektory systemowe \boldsymbol{u}_i mają zatem postać :

32)
$$\boldsymbol{u}_{i} = \begin{bmatrix} \boldsymbol{y}_{i} \\ \boldsymbol{x}_{i} \end{bmatrix}, i = 1, \dots, N, N = m, dim \boldsymbol{u}_{i} = m + n$$

Stąd otrzymuje się stwierdzenie:

Każdy obraz $I \dim(n \times 1)$ może być zrekonstruowany w systemie z rysunku 1 przy znajomości jego projekcji I_p :

$$(33) I_p = AI$$

Jakość rekonstrukcji \tilde{i} mierzona błędem MSE zależy od doboru parametru γ . Przykładową rekonstrukcję obrazu jako rozwiązanie problemu odwrotnego pokazano na rysunku 5.



Rys. 5. Rekonstrukcja obrazu w systemie z losowo wygenerowanymi wektorami systemowymi ($\gamma=0.1$)

Opis doświadczenia przedstawionego na rysunku 5.

Synteza systemu:

- Wygenerowano losowe wektory obrazu x_i , i = 1, ..., 2800, dim $x_i = 1096$.
- Wygenerowano losową macierz projekcji *A*, dim *A* = (2800 × 1096) (współczynniki macierzy *A* są losowe o rozkładzie normalnym).
- Wyznaczono projekcje y_i wektorów x_i :

$$y_i = Ax_i, i = 1, ..., 2800, \dim y_i = 2800$$

Wyznaczono wektory systemowe

$$\boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{y}_i \\ \boldsymbol{x}_i \end{bmatrix}$$
, $i = 1, \dots, 2800$, dim $\boldsymbol{u}_i = 4096$

Zgodnie z procedurą zapisaną równaniami (3)-(15) dokonano syntezy systemu z rysunku 1.

Rekonstrukcja obrazu

Analizie poddany był obraz o wymiarze (36×36 pikseli), który w procesie konkatenacji został przekształcony w wektor I, dim I = 1096. Projekcje uzyskano zgodnie z wzorem (33).

$$I_p = AI$$
, dim $I_p = 2800$

Na potrzeby rysunku wektor I_p przedstawiono w postaci macierzy odcieni szarości (50×56 pikseli).

Na wejście systemu wprowadzono wektorową projekcję I_p analizowanego obrazu. System rekurencyjnie zrekonstruował wektor \tilde{I} przypisany do oryginalnego obrazu. Obraz widoczny na rysunku 5 powstał w procesie odwrotnym do konkatenacji z wektora Ĩ.

Jakość rekonstrukcji była oceniana na podstawie błędu średniokwadratowego $MSE(I, \tilde{I})$ między oryginalnym obrazem a uzyskaną w systemie repliką. Wartości błędu średniokwadratowego w zależności od wartości parametru regularyzującego γ w równaniu (31) zamieszczono w tabeli 1.

Tabela 1. Zależność błędu średniokwadratowego rekonstrukcji obrazu od wartości parametru regularyzującego

Parametr regularyzujący γ	$MSE(I, \tilde{I})$
0.5	4.7·10 ⁻¹⁰
0.1	6.2·10 ⁻⁸
0.01	1.8·10 ⁻⁴
0.005	0.003
0.003	0.022
0.002	0.154
0.001	proces rozbieżny

Na rysunku 6 zobrazowano graficznie wpływ parametru γ na jakość rekonstrukcji obrazu w systemie uczenia maszynowego. Zaobserwowano pewna graniczną wartość parametru regularyzującego poniżej proces której rekonstrukcji jest rozbieżny.



 $\gamma = 0.005$

γ=0.003

Rys. 6. Wpływ parametru regularyzującego γ na jakość rekonstrukcji obrazu

Podsumowanie

Celem artykułu jest prezentacja rozwiązania problemu odwrotnego z wykorzystaniem autorskiego modelu uczenia Źródłem maszynowego. rozpatrywanego problemu

odwrotnego są równania liniowe, które są często stosowanymi modelami w rekonstrukcji obrazów. Szczególnie interesujący jest model uczenia maszynowego oparty na losowych wektorach systemowych. Model ten stwarza także przesłanki na jego zastosowanie w systemach łączności wykorzystujących technologię MIMO.

Autorzy: dr inż. Wiesław Citko, Uniwersytet Morski Gdynia, Wydział Elektryczny, ul. Morska 81-87, 81-225 Gdynia, E-mail: w.citko@we.umg.edu.pl; dr hab. inż. Wiesław Sieńko, Uniwersytet Morski Gdynia, Wydział Elektryczny, ul. Morska 81-87, 81-225 Gdynia, E-mail: w.sienko@we.umg.edu.pl.

LITERATURA

- [1] Pal S.K., Ghosh A., Kundu M.K. (Eds.), Soft Computing for Image Processing, In Studies in Fuzziness and Soft Computing, Physica-Verlang Heidelberg (2000), New York, NY, USA.
- [2] Huang Z., Ye S., McCann M.T., Ravishankar S., Model-based Reconstruction with Learning: From Unsupervised Supervised and Beyond, arXiv (2021), arXiv:2103.14528v1.
- Kaderuppan S. S., Wong W.W.L., Sharma A., Woo W.L., [3] Smart Nanoscopy: A Review of Computational Approaches to Achieve Super-Resolved Optical Microscopy, IEEE Access (2020), no.8, 214801-214831.
- [4] Ramanarayanan S., Murugesan B., Ram K., Sivaprakasam M., DC-WCNN: A Deep Cascade of Wavelet Based Convolutional Neural Networks for MR Image Reconstruction, In Proceedings of the IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3-7 April 2020.
- [5] Fessler A.J., Optimization Methods for Magnetic Resonance Image Reconstruction: Key Models and Optimization Algorithms, IEEE Signal Process. Mag., 37 (2020), 33-40.
- [6] Zheng H., Sherazi S.W.A., Son S.H., Lee J.Y., A Deep Convolutional Neural Network-Based Multi-Class Image Classification for Automatic Wafer Map Failure Recognition in Semiconductor Manufacturing, Appl. Sci., 11 (2021), 9769.
- [7] Szhou S.K., Greenspan H., Davatzikos C., Duncan J.S., Ginneken B., Madabhushi A., Prince J.L., Rueckert D., Summers R.M., A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies with Progress Highlights, and Future Promise Proc. IEEE, 109 (2021), 820-838.
- [8] Quan T.M., Nguyen-Duc T., Jeong W.K., Compressed Sensing MRI Reconstruction Using a Generative Adversarial Network with a Cyclic Loss. *IEEE Trans. Med. Imaging*, 37 (2018), 1488–1497.
- [9] Mardani M., Gong E., Cheng J.Y., Vasanawala S.S., Zaharchuk G., Xing, L., Pauly J.M., Deep Generative Adversarial Neural Networks for compressive sensing MRI, IEEE Trans. Med. Imaging, 38 (2019),167-179.
- [10]Liang D., Cheng J., Ke Z., Ying L., Deep Magnetic Resonance Image Reconstruction: Inverse Problems Meet Neural Networks. IEEE Signal Process. Mag., 37 (2020), 141-151.
- [11]Citko W., Sienko W., Hamiltonian and Q-Inspired Neural Network-Based Machine Learning, IEEE Access, (2020), no. 8, 220437-220449.
- [12]Citko W., Sienko W., Inpainted Image Reconstruction Using an Extended Hopfield Neural Network Based Machine Learning System, Sensors, 22 (2022), no. 3.
- [13]Citko W., Trzebiatowski A., Sienko W., Rekonstrukcja zdjęć twarzy z wykorzystaniem systemu uczenia maszynowego, Przegląd Elektrotechniczny, 97 (2021), nr.12, 191-194
- [14]Gilton D., Ongie G., Willett R., Deep Equilibrium Architectures for Inverse Problems arXiv in Imaging. (2021), arXiv:2102.07944v2.
- [15]Arridge S., Maass P., Oktem O., Schonlieb C., Solving Inverse Problems using Data-driven Models, Acta Numer., 28, (2019), 1-174
- [16]Ongie G., Jalal A., Metzler C.A., Baraniuk R.G., Dimakis A.G., Willett R., Deep Learning Techniques for Inverse Problems in Imaging, arXiv (2020), arXiv:205.06001.