

Paweł Wołoszyn*

Model sieci neuronowej zliczającej obiekty w obrazie**

1. Wprowadzenie

Czynności centralnego układu nerwowego człowieka związane z przetwarzaniem postrzeganego obrazu stanowią obszar zainteresowań wielu gałęzi biocybernetyki. Szereg różnorodnych metod i narzędzi wykorzystywanych jest do tworzenia modeli odzwierciedlających funkcjonowanie ośrodków wzrokowych i ich rolę w analizowaniu bodźców wizualnych. W niniejszej pracy opisywany jest model sieci neuronów przeprowadzającej zliczanie obiektów obecnych w prezentowanym sieci obrazie. Model ten odwołuje się do pewnych biologicznych własności czynnościowych neuronów, jak sumowanie i progowanie pobudzeń czy refrakcja bezwzględna i nosi w sobie elementy takich koncepcji, jak sztuczne sieci neuronowe [4, 5], sieci PCNN [2] czy automaty komórkowe [3]. Jako podłoże ideowe modelu przyjęty został agentowy system dynamiczny skonstruowany według założeń opisanych w [7], choć podobny model można zrealizować także za pomocą innych wymienionych podejść zachowując jego ogólne założenia.

Przedstawiony poniżej model może przywołać na myśl pewne podobieństwo do sieci PCNN (*Pulse Coupled Neural Networks*), złożony jest bowiem z neuronów posiadających określoną charakterystykę czasową generowanych impulsów, prezentowany sieci obraz wpływa na wartości progów pobudzenia zaś analizowana odpowiedź sieci ma postać zmiennego w czasie sygnału. Podobieństwo obejmuje także bliskie związki z biologicznym pierwowzorem, który usiłuje się naśladować [1]. Konstrukcja i zasada działania prezentowanego modelu jest jednak odmienna pod wieloma względami od koncepcji PCNN, między innymi sieć ma topologię warstwową, komórki nie wykazują cyklicznej, pulsacyjnej dynamiki a działanie sieci opiera się na jednokrotnym przejściu fali pobudzenia przez jedną z warstw, podczas którego pozostałe warstwy generują serie impulsów.

2. Budowa modelu

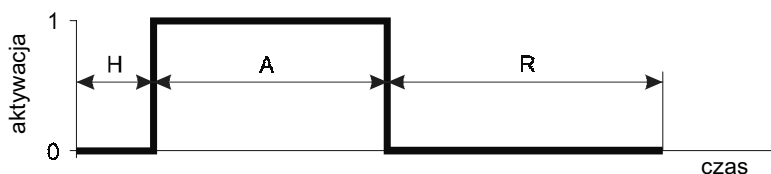
Modelowana sieć neuronów złożona jest z wielu komórek posiadających wspólną charakterystykę fizjologiczną opartą w pewnej mierze na biologicznym pierwowzorze komórki nerwowej. Podstawową czynnością neuronu jest przekazywanie bodźców z bio-

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie; pawel.woloszyn@agh.edu.pl

** Praca wykonana w ramach badań własnych – umowa AGH nr. 10.10.120.39

elektrycznego wejścia na wyjście, które przybiera postać potencjału czynnościowego. Ten fizjologiczny termin określa charakterystyczny przebieg zmian potencjału komórki pobudliwej w czasie, który składa się z kilku faz. Punktem początkowym potencjału czynnościowego jest pobudzenie wywołane dostatecznie dużym impulsem wejściowym zdolnym do przekroczenia progu pobudliwości.

Po wstępnej fazie utajonej aktywacji, podczas której komórka nie prezentuje jeszcze zmian potencjału, ale już dochodzi do wyzwolenia mechanizmu jego generowania, następuje zasadnicza faza potencjału czynnościowego, depolaryzacja, która dalej dla przejrzystości będzie określana jako aktywacja komórki. W modelu założono dla uproszczenia odmiennie niż w elektrobiologii oznaczenie potencjału błonowego komórki, a mianowicie potencjał w stanie spoczynku uznano za zerowy, natomiast wartość potencjału w fazie aktywacji przyjęto za równą 1. Faza aktywacji, podczas której potencjał komórki niezmiennie wynosi 1 i bodziec ten może być przekazywany do innych neuronów, trwa przez pewien czas, po którym następuje powrót potencjału do wartości zerowej i rozpoczyna się faza refrakcji. W fazie tej neuron jest niewrażliwy na bodźce, nawet jeśli przekraczają próg pobudliwości. Dopiero po upływie okresu refrakcji, która ma tu charakter refrakcji bezwzględnej, możliwe jest powstawanie kolejnych potencjałów czynnościowych. Przebieg potencjału czynnościowego przyjętego w modelu obrazuje rysunek 1.



Rys. 1. Przebieg potencjału czynnościowego w modelowanym neuronie.

H – faza utajonej aktywacji, A – faza aktywacji, R – faza refrakcji bezwzględnej

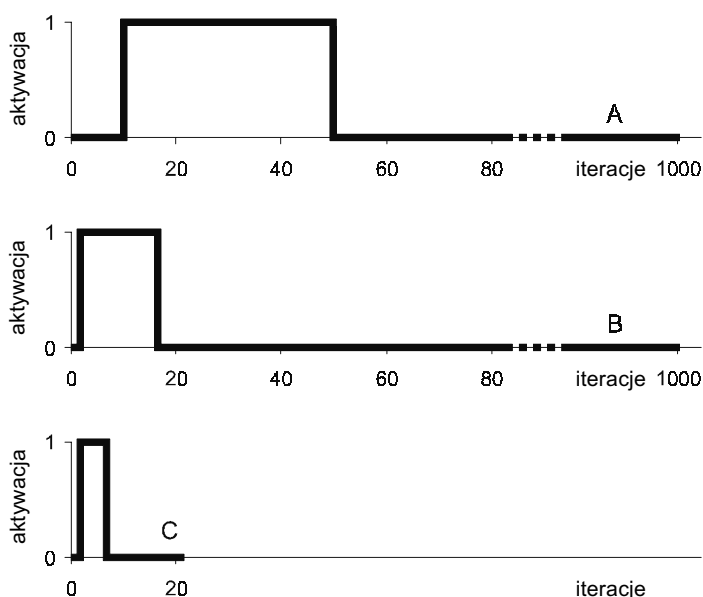
Do zbudowania sieci wykorzystano trzy populacje komórek: dystrybutorów, detektorów i transmiterów (o wzajemnych proporcjach liczebności 4:3:3) różniące się czasem trwania poszczególnych faz potencjału czynnościowego oraz progiem pobudliwości. Ich charakterystykę zawiera tabela 1.

Tabela 1
Populacje komórek występujących w modelowanej sieci

Typ	Fazy potencjału czynnościowego [iteracje]			Próg
	H	A	R	
Dystrybutor	10	40	1000	0,01
Detektor	1	15	1000	1000 lub 0,01*
Transmitter	1	5	15	0,01

* Próg zależy od zawartości obrazu prezentowanego w danym miejscu komórkom.

Komórki pełniące rolę dystrybutorów cechuje wolne przewodnictwo pobudzeń. Faza aktywacji trwa długo, natomiast silnie wydłużona faza refrakcji zapobiega nawracaniu i krążeniu pobudzeń w zamkniętych pętach. Detektory to neurony o szybszej reakcji na bodźce lecz również długim okresie refrakcji. Populacja transmitterów z kolei wykazuje najszybszą i najkrótszą odpowiedź na bodźce, jak również bardzo krótką refrakcję, dzięki czemu komórki te zdolne są szybko przewodzić pobudzenia następujące jedno po drugim w krótkich odstępach czasu (rys. 2).



Rys. 2. Proporcje czasowe faz potencjału czynnościowego w populacjach komórek modelu.
A – dystrybutor, B – detektor, C – transmitter

Aby komórka została aktywowana, musi ulec wpływowi bodźca przekraczającego próg pobudzenia, może to jednak nastąpić jako skumulowany efekt wpływu wielu słabszych bodźców. Neurony modelu, podobnie jak rzeczywiste komórki nerwowe, dokonują sumowania czasowego i przestrzennego docierających do nich sygnałów zgodnie z regułą

$$S_{t+1} = c \cdot S_t + s_t \quad (1)$$

gdzie:

- S_t – skumulowany wpływ bodźców w iteracji t ,
- s_t – sygnał wejściowy w iteracji t ,
- c – stała repolaryzacji.

Analogicznie do rzeczywistych neuronów komórki modelu dążą do przywrócenia swojego spoczynkowego potencjału 0, przez co skumulowany bodziec podprogowy z upływem czasu ulegnie zanikowi jeśli nie zostanie wywołany potencjał czynnościowy. W doświadczeniach na tym modelu przyjęto wartość $c = 0,99$.

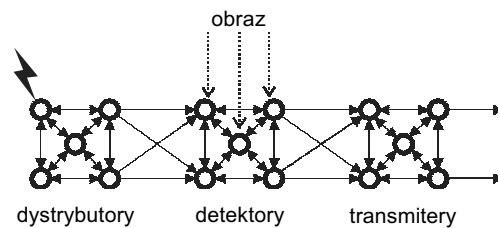
Sygnałem wejściowym dla neuronu są potencjały czynnościowe sąsiednich komórek, z którymi dany neuron posiada połączenia dendrytyczne, doprowadzające sygnał. Wagi tych połączeń zostały w modelu uzależnione od odległości między komórkami

$$s_i = \sum_{i=0}^n \frac{1}{1+d_i} a_i \quad (2)$$

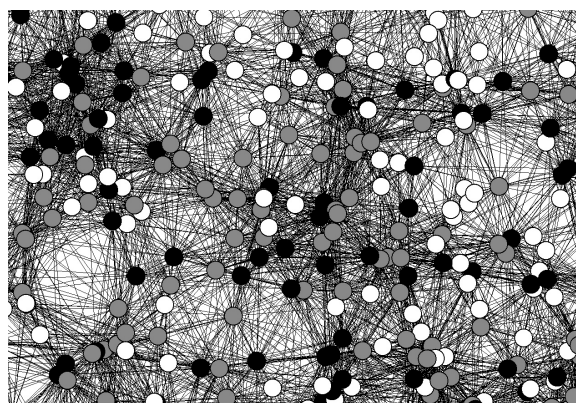
gdzie:

- n – liczba połączonych sąsiadów neuronu,
- a_i – aktywność i -tego sąsiada,
- d_i – odległość i -tego sąsiada od neuronu.

Sieć połączeń między komórkami tworzy się podczas inicjalizacji modelu w taki sposób, iż powstają trzy warstwy złożone z komórek tego samego typu. Połączenia doprowadzają sygnały od komórek danej warstwy do innych komórek tej samej warstwy oraz do komórek warstwy następnej (rys. 3). Rozkład przestrzenny komórek nie rozdziela jednak warstw od siebie, wszystkie typy komórek przemieszane są wzajemnie losowo w obrębie tego samego obszaru, zatem pojęcie warstwy ma tu wyłącznie charakter czynnościowy, związany z układem połączeń (rys. 4).



Rys. 3. Układ połączeń między komórkami warstw modelu



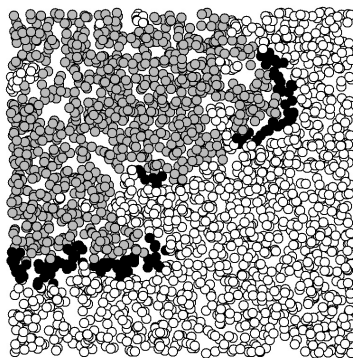
Rys. 4. Fragment sieci połączeń modelu. Różnymi kolorami oznaczono poszczególne populacje komórek

Z daną komórką połączone są jedynie komórki leżące w najbliższym sąsiedztwie, zarówno tego samego typu lub należące do następnej warstwy, wybrane według kryterium n najbliższych sąsiadów. Dla dystrybutorów przyjęta w modelu wartość n wynosi 4, dla detektorów 30, zaś dla transmitterów 50.

3. Działanie systemu

Symulacja pracy skonstruowanej sieci komórek rozpoczyna się od wygenerowania populacji komórek o losowym rozkładzie przestrzennym ograniczonym do umownego jednostkowego kwadratu na płaszczyźnie. Ustalone również zostają połączenia między komórkami zdeterminowane ich wzajemnymi odległościami. Po zakończeniu tych operacji sieci prezentowany jest obraz, który zyskuje swoją reprezentację w przestrzennym rozkładzie progów pobudzenia wśród komórek należących do detektorów w taki sposób, iż obszarom jasnym odpowiada próg o bardzo wysokiej wartości (1000), który czyni komórki praktycznie niepobudliwymi, zaś obszarom ciemnym odpowiada bardzo niski próg (0,01) prowadzący do łatwej pobudliwości komórek.

Po zaprezentowaniu obrazu rozpoczyna się zasadnicza, iteracyjna część symulacji, której punktem wyjścia jest sztuczna aktywacja jednego, arbitralnie wybranego neuronu warstwy dystrybutorów. Od niego począwszy, zaczyna rozprzestrzeniać się pobudzenie, które szerzy się wśród komórek dystrybutorów. Są to komórki o małej liczbie połączeń i wolnej reakcji, przez co front pobudzenia powoli dyfunduje po całym obszarze modelu z reguły w nieregularnej sekwencji (rys. 5). Długi okres refrakcji bezwzględnej zapobiega powtórnemu aktywowaniu dystrybutorów przez błędne dłuższe pobudzenia i powstaniu zamkniętych pętli.



Rys. 5. Rozprzestrzanie się pobudzenia wśród populacji dystrybutorów.

Kolor biały – komórki nieaktywne, czarny – aktywne, szary – komórki w fazie refrakcji

Gdy front aktywacji dotrze w pobliże komórek detektorów, których próg pobudzenia podczas prezentacji obrazu został ustalony na niskim poziomie, również i one ulegają aktywacji. Ich połączenia z sąsiednimi detektorami i bardzo szybka reakcja sprawiają, że pobu-

dzenie obejmuje w krótkim czasie wszystkie przyległe detektory, które wykazują pobudliwość, a więc odpowiadają spójnemu, ciemnemu obszarowi na prezentowanym sieci obrazie. Podobnie, jak w przypadku dystrybutorów, bardzo długi czas refrakcji zapobiega wielokrotnemu aktywowaniu tych samych grup detektorów podczas jednego przejścia frontu pobudzenia przez system. Dzięki temu pojedyncze obiekty w obrazie, będące oddzielnymi od otoczenia, spójnymi ciemnymi obszarami, będą wiązały się tylko z jednokrotnym pobudzeniem detektorów.

Pobudzenie detektorów przenosi się także na trzecią grupę komórek. Jak podkreśla to ich nazwa, są one przystosowane do transmitowania pobudzeń, a więc ich potencjał czynnościowy trwa krótko, szybko również kończy się okres refrakcji i w krótkim czasie komórki zyskują zdolność przewodzenia kolejnych pobudzeń. Liczba aktywnych komórek w populacji transmiterów, jest traktowana jako sygnał wyjściowy i rejestrowana podczas symulacji. Na podstawie przedstawionego zachowania systemu można spodziewać się, że liczba fal aktywacji odpowiada liczbie obiektów obecnych w obrazie prezentowanym systemowi.

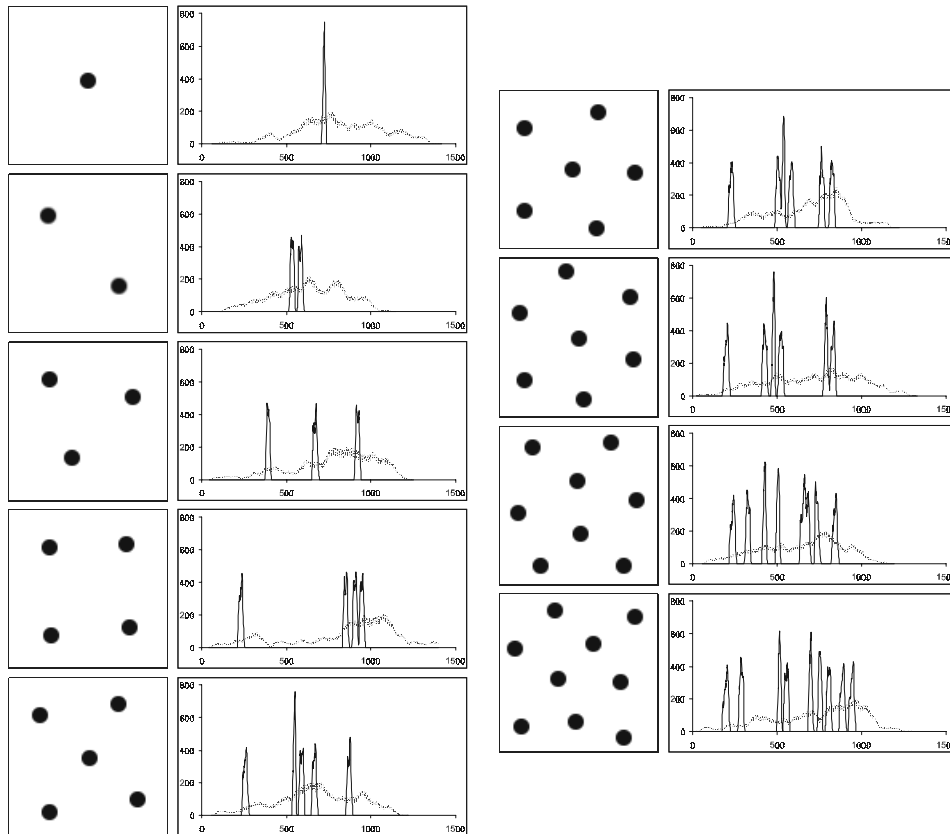
4. Rezultaty i wnioski

Przeprowadzone eksperymenty symulacyjne wskazują, że zachowanie systemu jest w dużej mierze zgodne z oczekiwaniami, a więc sieć generuje sygnał wyjściowy odpowiadający liczbie obiektów w obrazie reprezentowanej przez liczbę szczytów w przebiegu zmian liczby aktywnych komórek transmiterów. Przykładowe rezultaty eksperymentów dla obrazów o różnej liczbie obiektów przedstawiono na rysunku 6. Przebieg symulacji zależy od układu komórek (losowanego podczas inicjalizacji systemu), jak również od wyboru komórki startowej, od której rozpoczyna się propagacja pobudzenia dystrybutorów. Zebrane na rysunku przebiegi obrazują zachowanie systemu dla różnych obrazów, nie ilustrują zaś zmienności zachowań dla tego samego obrazu.

Sieć spełnia swoje zadanie w sposób obarczony pewnymi błędami. W zależności od przebiegu procesu rozprzestrzeniania się pobudzenia w warstwie dystrybutorów może dojść do sytuacji, w której dwa lub więcej skupisk detektorów ulegnie pobudzeniu jednocześnie lub prawie jednocześnie. W takiej sytuacji kilka szczytów w sygnale wyjściowym może nałożyć się na siebie uniemożliwiając ich rozróżnienie. Prawdopodobieństwo wystąpienia tego rodzaju niedokładności w zliczaniu obiektów jest tym większe, im więcej jest owych obiektów, im bliżej siebie się znajdują lub im słabiej są od siebie odseparowane. Aby zmniejszyć ryzyko takich błędów, można dążyć do takiego układu warstwy dystrybutorów, przy którym front pobudzenia, wyzwalaający kolejne szczyty, będzie możliwie wąski (świadczy o tym liczba aktywnych dystrybutorów, również zobrazowana na rys. 6), lecz będzie przemieszczał się w taki sposób, by nie pominąć żadnego obszaru sieci. Doświadczenia z różnymi strukturami warstwy dystrybutorów wykazały, że rzeczywiście takie rozumowanie jest słuszne.

Działanie systemu pod pewnymi względami przypomina zachowanie człowieka, którego postawiono przed zadaniem policzenia obiektów w obrazie bez użycia dodatkowych narzędzi, jak na przykład przyborów do pisania i zaznaczania. Im więcej jest tych obiektów lub im trudniej je oddzielić od siebie, tym więcej błędów popełni liczący. W zależności od

perspektywy widzenia obiekty dające się odseparować w jednym rzucie mogą ulec nałożeniu na siebie w innym prowadząc do błędnego wyniku zliczania. Oprócz niefortunnego układu przestrzennego sama gęstość upakowania obiektów, rosnąca wraz z ich liczbą, może przekroczyć zdolności człowieka do ich policzenia bez jawnego numerowania.



Rys. 6. Wyniki symulacji dla serii obrazów o różnej liczbie obiektów. Na wykresach przedstawiono zmiany liczby aktywnych transmiterów (linia ciągła) oraz dystrybutorów (linia przerywana) w czasie symulacji wyrażonym w iteracjach

Błędy popełniane w wyniku nałożenia się kilku pobudzeń wskazują na możliwość poddania sygnału wyjściowego dodatkowej analizie. Na przykład obserwując czas trwania skoków aktywności transmiterów można zauważyć, że w przypadku prawie równoczesnej detekcji kilku obiektów szerokość szczytu jest większa niż w przypadku pojedynczych obiektów. Interesujące wydaje się również bliższe zbadanie własności zaproponowanego modelu w kontekście neurofizjologii podobnych czynności realizowanych przez mózg człowieka, co może dostarczyć inspiracji do dalszego rozwoju podobnych modeli systemów przetwarzających obrazy.

Literatura

- [1] Eckhorn R., Reitboeck H.J., Arndt M., Dicke P.: *A neural network for feature linking via synchronous activity: results from cat visual cortex and from simulations*. [w:] Cotterill R.M.J. (ed.): *Models of Brain Function*, Cambridge University Press 1989, 255–272
- [2] Johnson J.L., Padgett M.L.: *PCNN Models and Applications*. *IEEE Transactions on Neural Networks*, t. 10, nr 3, 1999, 480–498
- [3] Kułakowski K.: *Automaty komórkowe*. Kraków, Akademia Górniczo-Hutnicza, Ośrodek Edukacji Niestacjonarnej 2000.
- [4] Mikrut Z., Tadeusiewicz R.: *Sieci neuronowe w przetwarzaniu i rozpoznawaniu obrazów*. [w:] Nałęcz M. (ed.): *Biocybernetyka i Inżynieria Biomedyczna 2000, Tom 6: Sieci Neuronowe*, Warszawa, PAN 2000, 459–494
- [5] Ślot K.: *Sieci neuronowe komórkowe: efektywne narzędzia przetwarzania informacji obrazowej*. *Zeszyty Naukowe Politechniki Łódzkiej*, nr 819, Wydawnictwo Politechniki Łódzkiej 1999, 188–197
- [6] Tadeusiewicz R.: *Sieci neuronowe*. Warszawa, PWN 1993
- [7] Wołoszyn P.: *System multiagentowy jako dyskretny system dynamiczny o skończonej przestrzeni stanów – ujęcie formalne*. *Półrocznik AGH Automatyka*, t. 8, z. 3, 2004, 293–302
- [8] Wołoszyn P.: *Podobieństwo strukturalne i funkcjonalne multiagentowych systemów chaotycznych i systemów biologicznych*. „Sztuczna Inteligencja w Inżynierii Biomedycznej”, materiały konferencyjne, Kraków 2004