

ИССЛЕДОВАНИЕ АРХИТЕКТУРЫ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ЗАДАЧИ РАСПОЗНАВАНИЯ ОБРАЗОВ

Землевский А.Д. Email: Zemlevskiy630@scientifictext.ru

Землевский Андрей Дмитриевич – бакалавр техники и технологий,
кафедра вычислительной техники,
Тульский государственный университет, г. Тула

Аннотация: в статье анализируются архитектурные составляющие сверточных нейронных сетей, а также осуществляется выбор наиболее подходящей архитектуры для реализации на последовательных машинах. Уделено внимание схемам взаимодействия слоев, устройству блока классификатора и блока выделения признаков. Производится разделение сверточных нейронных сетей по типу комбинирования слоев на группы. Также приводятся основные факторы, которые следует учитывать при подборе архитектуры сверточной нейронной сети для задачи распознавания образов.

Ключевые слова: нейронные сети, сверточные нейронные сети, архитектура сверточных сетей, распознавание образов.

RESEARCH OF ARCHITECTURE CONVOLUTIONAL NEURAL NETWORKS FOR THE PROBLEM OF PATTERN RECOGNITION Zemlevskiy A.D.

Zemlevskiy Andrey Dmitrievich – Bachelor of Engineering and Technology,
COMPUTER ENGINEERING DEPARTMENT,
TULA STATE UNIVERSITY, TULA

Abstract: the article analyzes the architectural components of convolutional neural networks, and also selects the most suitable architecture for implementation on serial machines. Attention is paid to the interaction schemes of layers, the device of the classifier block and the feature selection block. Separation of convolutional neural networks by the type of combining layers into groups is made. Also, the main factors that should be considered when selecting the architecture of a convolutional neural network for the problem of pattern recognition are given.

Keywords: neural networks, convolutional neural networks, the architecture of convolutional neural networks, pattern recognition.

УДК 004.932.72'1

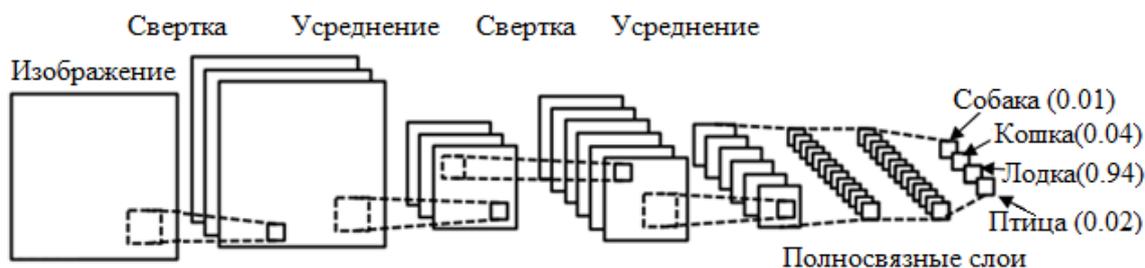


Рис. 0. Структура сверточной нейронной сети

Сверточная модель нейронной сети является широко модернизированным многослойным перцептроном, заточенным на получение наилучших результатов при распознавании изображений. Данная модель состоит из двух блоков: выделения признаков и классификатора.

Блок, отвечающий за выделение признаков, состоит из чередующихся слоев свертки и усреднения, данная концепция с чередованием сложных и простых клеток также используется в неокогнитроне. Каждый нейрон воспринимает сигнал из предыдущего слоя через поле восприимчивости, которая представляет собой матрицу смежных нейронов предыдущего слоя связанных с нейроном текущего слоя настраиваемыми весами. В процессе обучения данные слои формируют карты абстрактных признаков для каждого класса.

Блок классификации обычно представляется многослойным перцептроном, реже другими сетями, такими как SVM или RBF-сеть.

В среднем точность сетей данной модели превосходит обычные модели ИНС на 10-15% [1, 2]. Так же СНС является ключевой моделью технологии Deep Learning [3, 4]. Основной причиной успеха данной

модели стала концепция общих весов, поэтому несмотря порой на огромный размер сеть имеет небольшое количество настраиваемых весов по сравнению с неокогнитроном. Однако, имеются вариации [5] данной модели, концептуально схожие с неокогнитроном, из-за отказа в данных моделях от концепции разделения весов. Главное преимущество моделей СНС в том, что они могут относительно быстро работать и обучаться на последовательных машинах.

Сверточная нейронная сеть состоит из двух блоков: первый из них отвечает за выделение признаков, а второй за их классификацию, данное разделение наглядно продемонстрировано на рисунке 2.

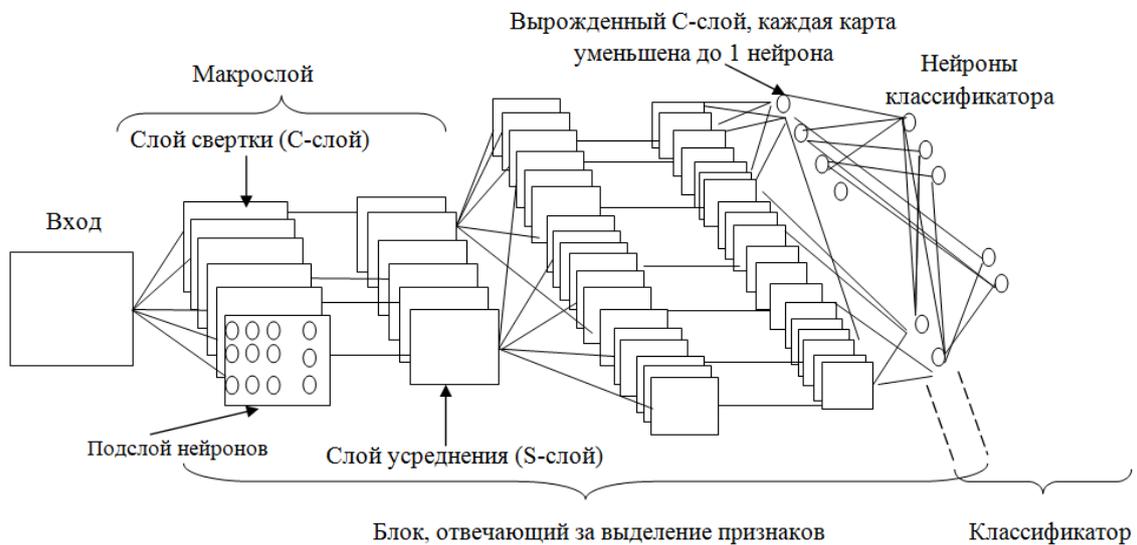


Рис. 2. Разделение структуры СНС на два логических блока

Структура блока выделения признаков состоит из вариативного количества слоев. Каждый слой – это группировка подслоев, которые обрабатывают входной сигнал с учетом топологии изображения. Архитектура СНС подразумевает два вида слоев: сверточный (С-слой) и усредняющий (S- слой). Данные слои обычно чередуются [6], и каждая пара таких слоев называется макрослоем. Каждый нейрон в блоке выделения признаков имеет свою область восприимчивости (поле охвата) [7, 8, 9, 10], которая представляет собой квадратную матрицу, включающую в себя нейроны передающие выходные сигналы нейрону следующего слоя. Описанная схема взаимодействия наглядно представлена на рисунке 3.

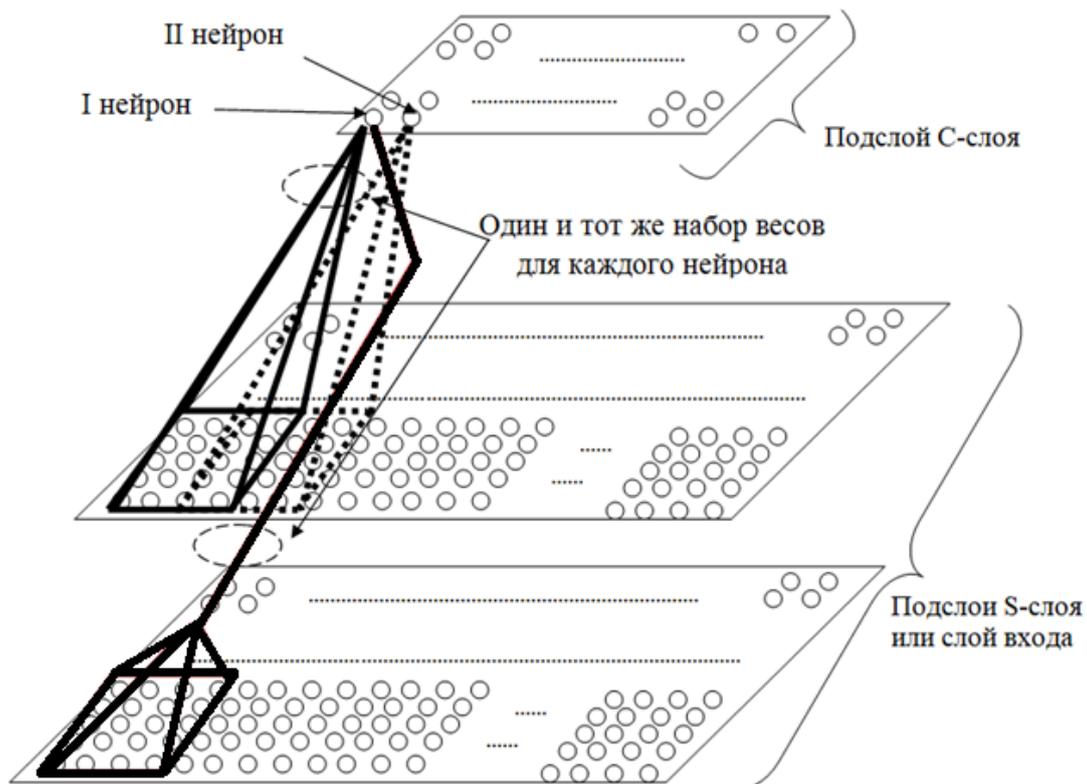


Рис. 3. Схема взаимодействия нейрона текущего слоя с предыдущим

Каждый подслой С – слоя может быть связан с произвольным количеством подслоев предыдущего S – слоя или же со слоем входа. Каждый нейрон С-подслоя имеет поле охвата идентично позиционированное на каждом из предшествующих подслоев. На рисунке указанные поля выделены сплошными линиями разных цветов размером 4*4. Получается, что каждый нейрон С - подслоя должен иметь 33 параметра с учетом параметра смещения, но это не совсем так потому что главная особенность сверточного подслоя - это общие веса поэтому для каждого нейрона используется общий набор весов. Общие веса позволяют сети реагировать на некоторый признак в любом месте подслоя (карты).

Пунктиром на рисунке 3 обозначено поле охвата следующего нейрона, также видно, что происходит пересечение текущего поля и предыдущего, данное смещение можно уменьшать для создания более подробной карты признаков. Поля охвата каждого нейрона и общие веса являются predetermined information, благодаря которой происходит выделение карты признаков [11].

Далее ниже на рисунке представлена схема нейрона С-подслоя.

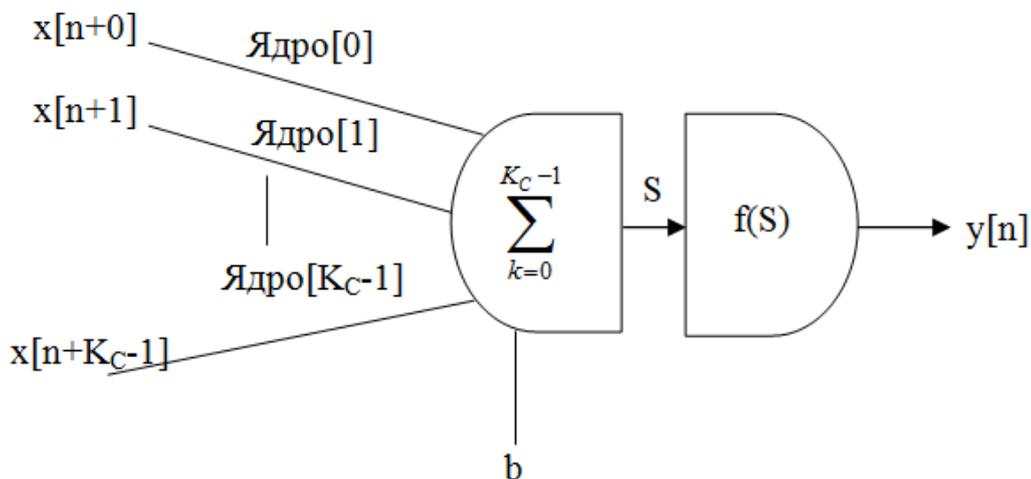


Рис. 4. Общая схема нейрона С-подслоя

K_c - общее количество нейронов, входящих в поле охвата n -го нейрона C -подслоя, Ядро $[k]$ – настраиваемые веса нейрона, b – смещение n -го нейрона, причем b и Ядро $[k]$ – одни и те же для нейронов всего C -подслоя, $x[n+k]$ – входной сигнал для n -го нейрона C -подслоя, $k = 0, K_c - 1$.

Взвешенная сумма входного сигнала $x[n+k]$ и параметров Ядро $[k]$ равна:

$$S = b + \sum_{k=0}^{K_c-1} \text{Ядро}_k * x_{n+k} \quad (1)$$

Выходной сигнал нейрона определяется функцией активации, которая принимает на вход взвешенную сумму S .

$$y = f(S) \quad (2)$$

Использование в качестве функции активации гиперболического тангенса является хорошим выбором если в качестве алгоритма обучения СНС используется алгоритм обратного распространения так как данная функция не линейная и хорошо дифференцируется.

Любой нейрон C -подслоя можно назвать детектором некоторого признака, выделяемого во время обучения. Каждый нейрон сверточного подслоя проводит свертку сигнала из поля охвата со своими весами, которые также используются для получения реакции других нейронов подслоя. Результат реакции свидетельствует о степени присутствия данного признака в области, охватываемой текущим нейроном. Свертка входного сигнала с общими настраиваемыми параметрами является аналогом фильтрации [12] изображения поэтому изображение, формируемое на подслое C -слоя, будет отфильтрованным изображением относительно настраиваемых весов.

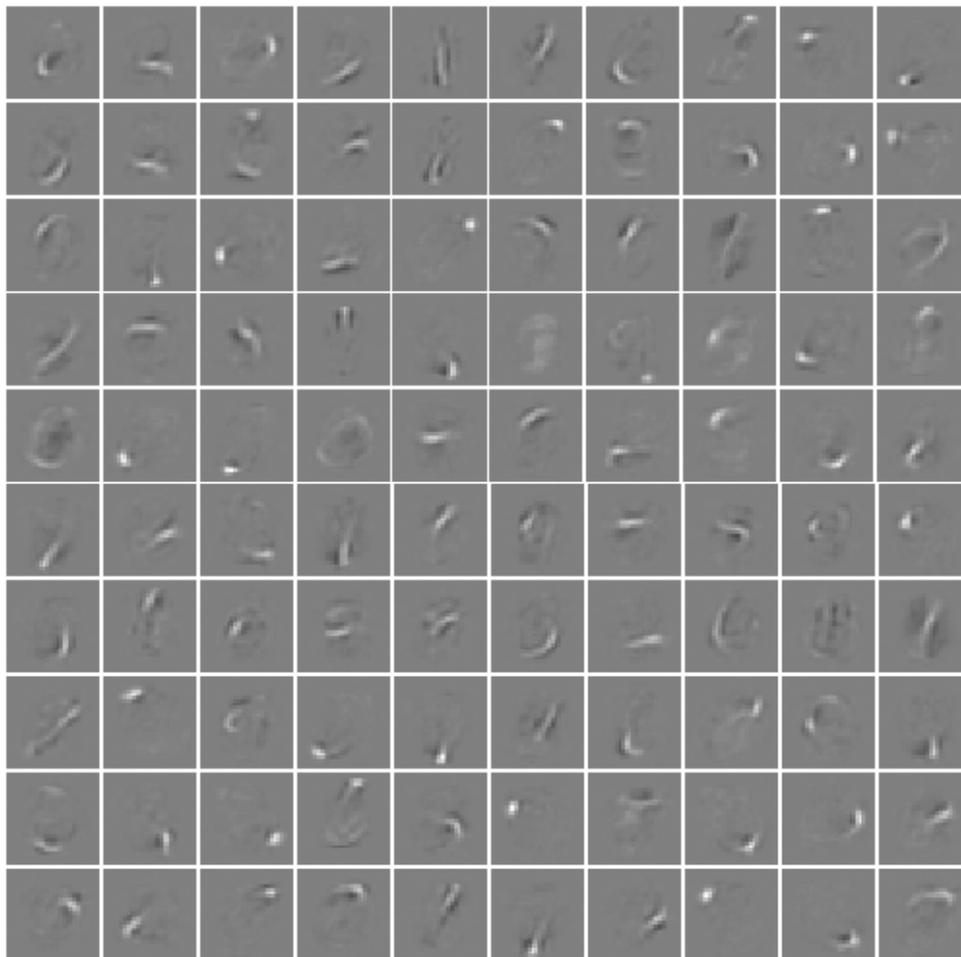


Рис. 5. Пример карты признаков

Далее представлена схема взаимодействия нейронов усредняющего подслоя с нейронами сверточного подслоя (рисунок 6).

Главная роль усредняющего подслоя в блоке выделения признаков – это облегчение выделения абстрактных признаков за счет уменьшения размерности отфильтрованного изображения, полученного на предыдущем сверточном подслое.

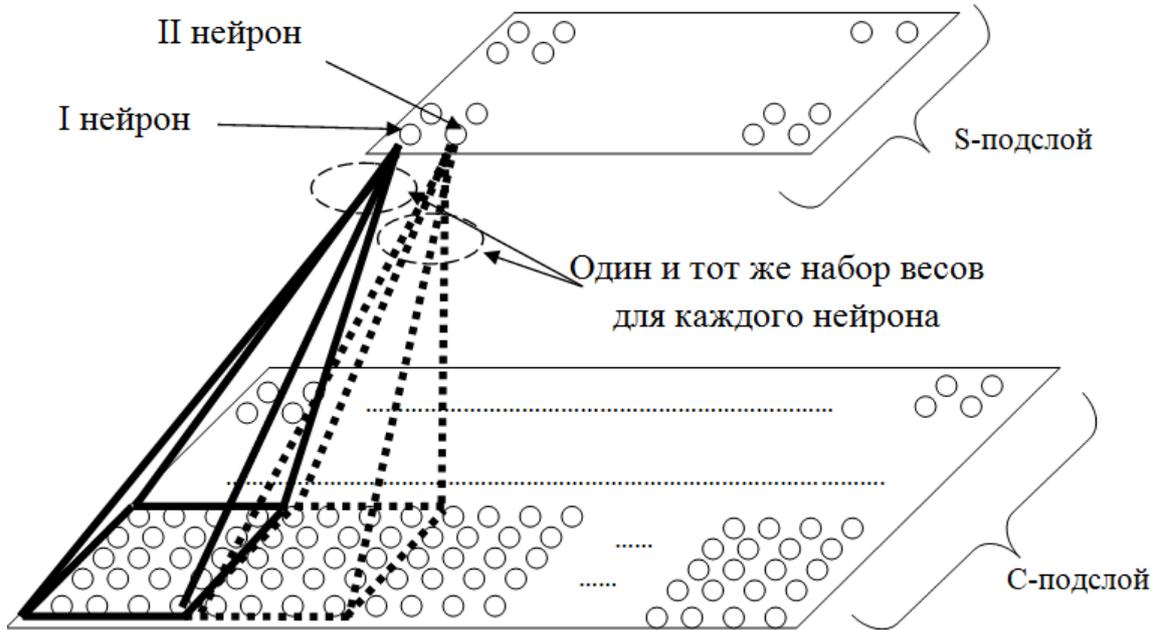


Рис. 6. Схема взаимодействия S-подслоя с предыдущим C-подслоем

Из рисунка 1.10 видно, что поля, охватываемые нейронами S-подслоя в отличие от C-подслоя не пересекаются. Как и в C-подслое настраиваемые параметры в S-подслое так же общие, но их количество всегда равно двум и совсем не зависит от количества элементов, охватываемых каждым нейроном S-подслоя.

Схема нейрона усредняющего подслоя представлена ниже на рисунке 1.8

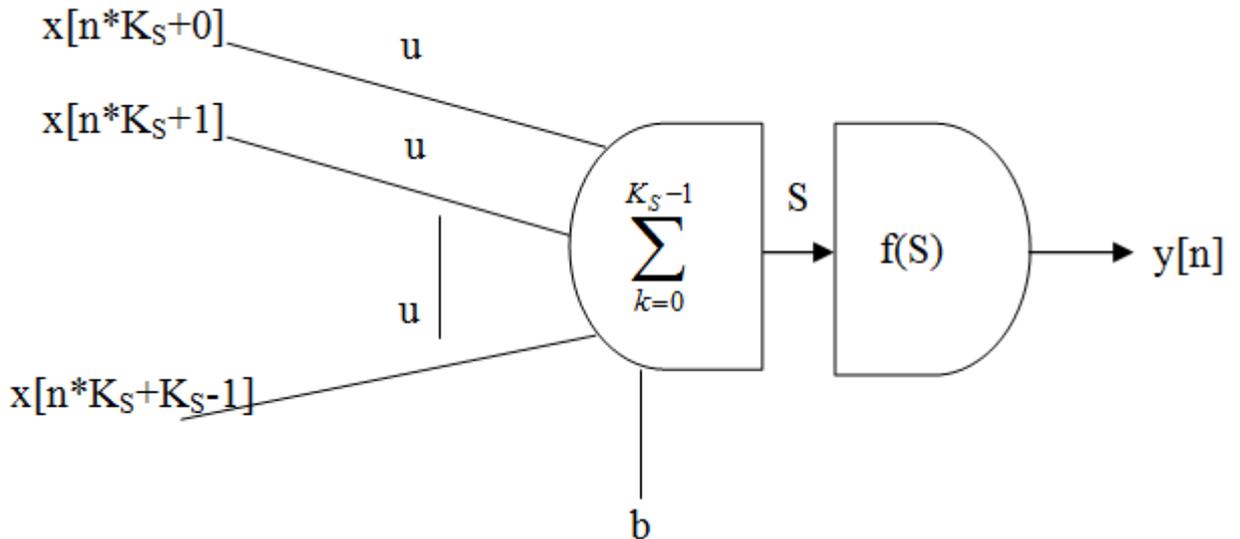


Рис. 7. Схема нейрона усредняющего подслоя

Из рисунка 7 видно, что все настраиваемые весовые параметры нейрона участвующие в свертке равны, второй параметр – это смещение b . Так же как и в сверточном подслое, указанные веса являются общими для каждого нейрона усредняющего подслоя. Таким образом, учитывая, что области охвата нейронов не

пересекаются, взвешенная сумма для n-го нейрона S-подслоя определяется как:

$$S = b + u * \sum_{k=0}^{K_s-1} x_n * K_s + k, \quad (3)$$

где K_s - количество нейронов, находящихся в поле охвата n-го нейрона S-подслоя.

Также из рисунка 1.8 видно, что размерность карт признаков от слоя к слою постепенно снижается, а их количество растет, дабы компенсировать потерю точности, в итоге давая на выходе блока выявления признаков множество карт единичной размерности.

Второй блок СНС – классификатор признаков, выявленных первым блоком, представленный двухслойным персептроном.

Карты признаков единичной размерности, полученные на выходе блока выявления, подаются на вход блоку классификации. Общая схема нейрона блока классификации представлена на рисунке 1.9.

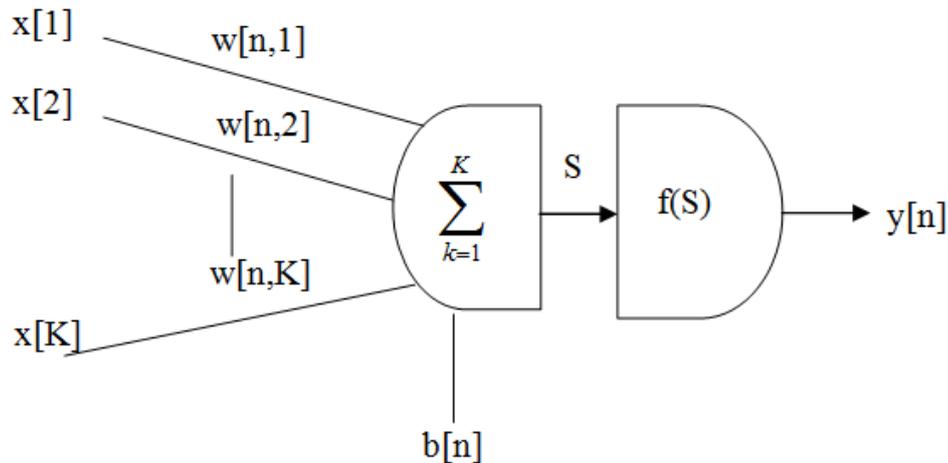


Рис. 8. Схема нейрона слоя классификации

Взвешенная сумма слоя классификации определяется как:

$$S = b_n + \sum_{k=1}^k x_k * w_{n,k}, \quad (4)$$

где b_n - смещение, уникальное для каждого нейрона, $x[k]$ – входной сигнал, $w[n,k]$ – настраиваемые параметры n-го нейрона, являющиеся уникальными для каждого нейрона, K – размер входного сигнала для слоя классификации.

В качестве функции активации используется для нейронов блока классификации была выбрана классическая для данной модели сети функция гиперболического тангенса:

$$f(S) = 1.7159 * \tanh\left(\frac{2}{3} * p\right), \quad (5)$$

СНС с такой функцией показывает наилучшие результаты по показателям ошибки распознавания, а также скорости сходимости ошибки обучения.

На основе анализа литературы [13, 14, 15, 16, 17, 18, 19, 20, 21, 22], описывающей моделирование и создание СНС, следует выделить два фактора, которые влияют на выбор архитектуры СНС.

Первый фактор – это тип обучения, которое может производиться как с учителем, так и без.

Обучение с учителем подразумевает наличие меченной выборки данных. Пример архитектуры такой сети представлен на рисунке 1, где блок классификатора представлен многослойным персептроном. Главные плюсы, получаемые от выбора данного типа обучения, - простота самого процесса обучения, а также архитектуры классификатора и вытекающая из этого точность обобщения. Главный минус – сложность создания репрезентативной выборки. Для систем машинного зрения, ориентированных на небольшое количество классов (<100), лучше всего подходит вариант обучения с учителем.

Обучение без учителя производится с помощью нейронных сетей со встроенной ассоциативной памятью, которые учатся отображать вход на выход. Основным плюсом таких сетей - отсутствие надобности формирования выборки для обучения, из чего вытекает и главный минус данного типа

обучения - это несбалансированность данных, подающихся на вход, что в итоге дает небольшую точность распознавания ~ 80%. Такой тип обучения логичнее применять для систем, требующих способности к распознаванию огромного количества классов, так как данные для обучения подобной сети можно подобрать в огромном количестве, потеряв при этом в качестве.

Второй фактор – подбор типов слоев для блока выделения признаков.

Разделим архитектуры СНС по типу комбинирования в них слоев разных видов на три группы:

1. Простейшая, использующая только слои свертки, где уменьшение размерности карт признаков происходит за счет уменьшения зона пересечения полей охвата нейронов. Хотя данная архитектура очень проста, но не практична так как добиться приличных результатов обобщения при распознавании невозможно без алгоритмов направленных на создание искажений. Такая архитектура применяется достаточно редко.

2. Классическая, включающая макрослой, состоящие из слоя свертки и усреднения. Все еще простая для последовательной реализации архитектура, использующая слои усреднения для уменьшения размерности карт признаков. Усреднение может осуществляться по разным алгоритмам по-разному влияя на обобщающую способность сети.

3. Усложненные, использующие множество типов слоев и их комбинации. В данной архитектуре могут использоваться слои для «выпрямления» нелинейности, а также слой локальной нормализации контраста. Данные слои позволяют добиться улучшения точности распознавания, но затрудняют последовательную реализацию данной архитектуры.

Список литературы / References

1. *Галушкин А.И.* Нейрокомпьютеры для обработки изображений [Текст] / А.И. Галушкин, Н.С. Томашевич, Е.И. Рябцев // Нейрокомпьютеры в прикладных задачах обработки изображений. Кн. 25. 2007. С. 74–109.
2. *Krizhevsky A.* ImageNet classification with deep convolutional neural networks [Text] / A. Krizhevsky, I. Sutskever, G.E. Hinton // Advances in Neural Information Processing Systems 25, 2012. Pp. 1106-1114.
3. *Bengio Y.* Learning deep architectures for AI [Text] / Y. Bengio // Foundations and Trends in Machine Learning. Vol. 2. Issue 1, 2009. Pp. 1-127.
4. *Bengio Y.* Scaling learning algorithms towards AI [Text] / Y. Bengio, Y. LeCun // in Large Scale Kernel Machines, MIT Press, 2007.
5. *Le Q.* Tiled convolutional neural networks [Text] / Q. Le, J. Ngiam, Z. Chen, D. Chia, P. Koh, A.Y. Ng // NIPS, 2010.
6. *Ranzato M.A.* What is the best multi-stage architecture for object recognition? [Text] / M.A. Ranzato, K. Jarrett, K. Kavukcuoglu, Y. LeCun // In ICCV, 2009.
7. *Соколов Е.Н.* Нейроинтеллект: от нейрона к нейрокомпьютеру [Текст] / Е.Н. Соколов, Г.Г. Вайткявичюс // М.: Наука, 1989. 238 с.
8. *Шевелёв И.А.* Нейроны-детекторы зрительной коры [Текст] / И.А. Шевелёв // М.: Наука, 2010. 183 с.
9. *Eysel U.T.* Pharmacological studies on receptive field architecture [Text] / U.T. Eysel // N.Y.: Acad. Press, 2002. Pp. 427-470.
10. *LeCun Y.* Backpropagation applied to handwritten zip code recognition [Text] / Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, L.D. Jackel // Neural Computation. Vol. 1, № 4, 1989. Pp. 541-551.
11. *Хайкин С.* Нейронные сети: полный курс, 2-е издание. Пер. с англ. [Текст] / С. Хайкин // М.: Издательский дом «Вильямс», 2008. 1104 с.: ил. Парал. тит. англ.
12. *Гонсалес Р.* Цифровая обработка изображений. Издание 3-е, исправленное и дополненное [Текст] / Р. Гонсалес, Р. Вудс // М.: Техносфера, 2012. 1104 с.
13. *Kavukcuoglu K.* Learning Feature Hierarchies for Object Recognition [Text] / K. Kavukcuoglu // PhD diss. New York University, January 2011.
14. *Kavukcuoglu K.* Learning invariant features through topographic filter maps [Text] / K. Kavukcuoglu, K. Ranzato, M. Fergus, Y. LeCun // In CVPR'09, IEEE, 2009.
15. *Larochelle H.* An empirical evaluation of deep architectures on problems with many factors of variation [Text] / H. Larochelle, D. Erhan, A. Courville, J. Bergstra, Y. Bengio // In Twenty-fourth International Conference on Machine Learning (ICML, 2007).
16. *Neal R.M.* Connectionist learning of belief networks [Text] / R.M. Neal // PhD thesis, Department of Computer Science. University of Toronto, 1994.
17. *Ackley D.H.* A learning algorithm for Boltzmann machines [Text] / D.H. Ackley, G.E. Hinton, T.J. Sejnowski // Cognitive Science. Vol. 9, 1985. Pp. 147-169.
18. *Ranzato M.* Unsupervised learning of invariant feature hierarchies with applications to object recognition

- [Text] / M. Ranzato, F. Huang, Y. Boureau, Y. LeCun // in Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR'07). IEEE Press, 2007.
19. *Le Q.* Building high-level features using large scale unsupervised learning [Text] / Q. Le, A. Ranzato, R. Monga, M. Devin, G. Corrado, K. Chen, J. Dean // Acoustics, Speech and Signal Processing (ICASSP), 2013.
 20. *Farabet C.* Hardware accelerated convolutional neural networks for synthetic vision systems [Text] / C. Farabet, B. Martini, P. Akselrod, S. Talay, Y. LeCun, E. Culurciello // In International Symposium in Circuits and Systems (ISCAS'10). IEEE. Paris, May 2010.
 21. *Ciresan D.* Multicolumn Deep Neural Networks for Image Classification [Text] / D. Ciresan, U. Meier, J. Schmidhuber // In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). CVPR '12. Pp. 3642-3649. Washington. DC. USA, 2012. IEEE Computer Society.
 22. *Червяков Н.И.* Нейрокомпьютеры в остаточных классах. Кн. 11. [Текст] / Н.И. Червяков, П.А. Сахнюк, А.В. Шапошников, А.Н. Макоха // М.: Радиотехника, 2003. 272 с.: ил.