

Исследование эффективности использования свойства нейропластичности в свёрточных сетях

Г.А. Альгашев¹, О.П. Солдатова¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. При обучении нейронных сетей часто можно столкнуться со следующими проблемами: малая обучающая выборка, долгое время обучения нейронной сети, проблема исчезающего градиента. Эти проблемы долгое время были препятствием для обучения действительно глубоких нейронных сетей. Для решения поставленных проблем получил развитие метод переноса знаний, основанный на свойстве нейропластичности нейронных сетей. Идея свойства нейропластичности в свёрточных сетях состоит в том, чтобы на хорошо обученную сеть подать новый класс изображений, на которых сеть не обучалась раньше, и продолжить обучение сети с новым набором данных, таким образом веса в верхних слоях сети изменятся незначительно, а в глубоких слоях будут активно изменяться веса. В статье свойство нейропластичности исследовано на примере решения задачи классификации маркированных клеток крови. Для исследования обучаются несколько моделей нейронных сетей, с использованием метода нормализованной инициализации весов, метода переноса знаний и метода смешанной инициализации весов. С увеличением количества слоёв, в которых веса инициализируются с использованием метода переноса знаний, точность моделей увеличивается. Лучший результат показала модель, у которой все слои были инициализированы с использованием свойства нейропластичности.

1. Введение

Одной из классических задач, которую решают с помощью нейронных сетей является задача классификации изображений. Для этих целей чаще всего используют свёрточные нейронные сети.

Во время обучения сети программист чаще всего сталкивается со следующими проблемами:

- малая обучающая выборка, в результате чего не удаётся обучить нейронную сеть до требуемой погрешности распознавания;
- долгое время обучение сети, в результате чего поиск наилучшего варианта архитектуры сети сильно усложняется.

Для решения первой проблемы чаще всего используют различные алгоритмы искажения обучающей выборки, чтобы повысить количество изображений, на которых сеть будет обучаться. А для решения проблемы времени обучения используют механизмы распараллеливания вычислений.

Но, не смотря на обилие алгоритмов и методов для решения выше указанных проблем, справиться с ними в полной мере не удаётся. Поэтому исследования в этой области в настоящее время активно ведутся и будут актуальными в ближайшем будущем.

В настоящее время большую популярность получает метод переноса знаний для обучения нейронных сетей [1]. Он основан на свойстве нейропластичности, который присущ нейронам головного мозга [2]. Его идея заключается в использовании уже полученного опыта для решения новой задачи. В нейронных сетях «опытом» являются веса хорошо обученной нейронной сети, которые затем будут использоваться для обучения новой модели. Особенностью данного метода является то, что его можно использовать вместе с другими алгоритмами обучения нейронных сетей, тем самым понижая погрешность распознавания изображений.

1.1. Описание свойства нейропластичности

Под нейропластичностью чаще всего понимают свойство головного мозга изменяться под действием некоторого опыта, а также восстанавливать утраченные связи после травмы [2]. Эти изменения включают в себя создание новых синаптических связей и создание новых нейронов. До 70-х годов 20 века учёные считали, что человеческий мозг не меняется с детства и остаётся все время статичным [3]. Но новые открытия, сделанные за последние десятилетия, показали, что эти утверждение не верны. Мозг адаптивен, и этому свойству дали название нейропластичность. Процесс обучения заключается в возникновении, укреплении, ослаблении или разрушении связей между нейронами головного мозга. Также большое влияние оказывает процесс нейрогенеза – создание в головном мозге новых нейронов [4].

Свойство нейропластичности можно описать следующим образом. В процессе выполнения какого-либо действия человеком, в его головном мозге начинают задействоваться определённые нейроны, и чем чаще это действие выполняется, тем больше укрепляются связи между нейронами. Если же действия редко выполняются человеком, то связи между активизирующимися нейронами ослабевают, и в результате могут совсем разрушиться. Если же человек выполняет действие, которое до этого он не совершал, то между активизирующимися нейронами будут создаваться новые связи, которых до этого не было в головном мозге.

1.2. Нейропластичность в нейронных сетях

Для решения задачи классификации изображений чаще всего используют свёрточные нейронные сети [5]. Классическая свёрточная нейронная сеть состоит из слоёв свёртки и слоёв подвыборки, которые чередуются между собой, и выходного слоя (рисунок 1). Сеть является однонаправленной и многослойной.

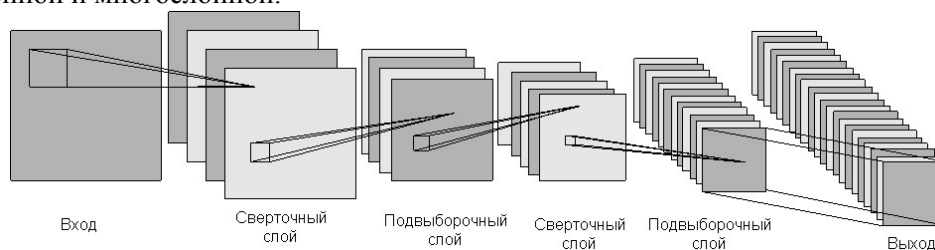


Рисунок 1. Архитектура свёрточной нейронной сети.

В 2013 году Метью Зейлер и Роб Фергус написали статью «Visualizing and Understanding Convolution Networks», в которой они поставили себе цель визуализировать признаки, которые извлекает свёрточная нейронная сеть на каждом своём слое [6]. Результаты их работы можно увидеть на рисунке 2.

На верхних слоях нейронная сеть извлекает низкоуровневые признаки изображений, а на глубоких слоях сеть извлекает высокоуровневые признаки изображений. В результате исследований так же был сделан вывод, что на первых слоях нейронная сеть извлекает фильтры Габора, которые широко используются в компьютерном зрении для распознавания границ объектов на изображении [7].

Дальнейшие исследования свёрточных нейронных сетей на различных наборах изображений показали, что сети на верхних слоях извлекают очень похожие между собой низкоуровневые

признаки, а на глубоких слоях высокоуровневые признаки уже сильно отличаются [1]. Таким образом, во время обучения свёрточных нейронных сетей, сеть каждый раз извлекает очень похожие между собой низкоуровневые признаки.

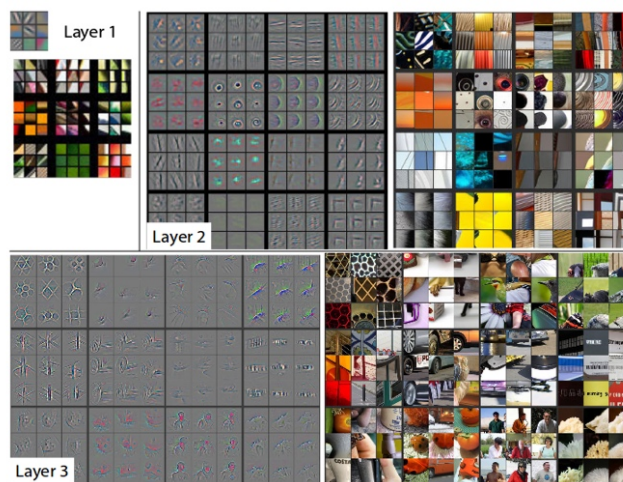


Рисунок 2. Визуализация свёрточных слоёв нейронной сети.

В свёрточных нейронных сетях чаще всего для вычисления градиента используется метод обратного распространения ошибки [8]. В результате использования этого метода возникает проблема исчезающего градиента, из-за которого передние слои нейронной сети обучаются очень медленно [9].

Исходя из выше сказанного, можно сделать вывод, что во время обучения свёрточная нейронная сеть каждый раз тратит большое количество времени на обучение верхних слоёв, которые в итоге извлекают похожие низкоуровневые признаки. Для решения этой проблемы и применяют свойство нейропластичности.

Идея свойства нейропластичности в нейронных сетях состоит в том, чтобы на хорошо обученную нейронную сеть подать новый класс изображений, на которых сеть не обучалась раньше, и продолжить обучение сети с новым набором данных. Для решения задачи классификации нового набора изображений вероятнее всего придётся изменить выходной слой нейронной сети. Во время обучения верхние слои будут изменяться незначительно, т.к. они уже умеют хорошо извлекать низкоуровневые признаки, а глубокие слои будут интенсивно менять свои веса, извлекая высокоуровневые признаки нового обучающего набора изображений. Таким образом, нам не придётся обучать нейронную сеть с нуля, что было бы при случайной инициализации всех весов.

1.3. Постановка задачи нейропластичности

Исследования свойства нейропластичности в нейронных сетях начались в 2009 году со статьи Лизы Торрей и Джуда Шавлика «Transfer Learning» [1]. В ней авторы выделили три основные цели использования нейропластичности, или как его ещё называют метод переноса знаний (рисунок 3):

- быстрый старт – улучшение качества обучения уже на начальных итерациях;
- большой наклон – ускорение сходимости алгоритмов обучения;
- большая асимптота – улучшение верхней достижимой границы качества обучения.

Для того, чтобы использовать свойство нейропластичности в нейронных сетях необходимо иметь эталонную хорошо обученную сеть, на основе которой будут решать новую задачу. Эталонная сеть может быть обучена на любых данных, главное, чтобы у неё была как можно меньше погрешность решения задачи. В процессе изменения сети, для решения новой задачи происходит подмена обучающего набора данных на новый, на котором эталонная модель будет дообучаться. В результате будет получена новая модель нейронной сети, которая будет решать уже новую задачу.

Использование эталонной модели необходимо для уменьшения временных затрат на создание и обучение нейронной сети, т.к. не придётся подбирать для неё архитектуру (количество слоёв, количество нейронов в каждом слое и т.д.), а веса нейронов не будут сильно изменяться, т.к. верхние слои уже хорошо обучены.

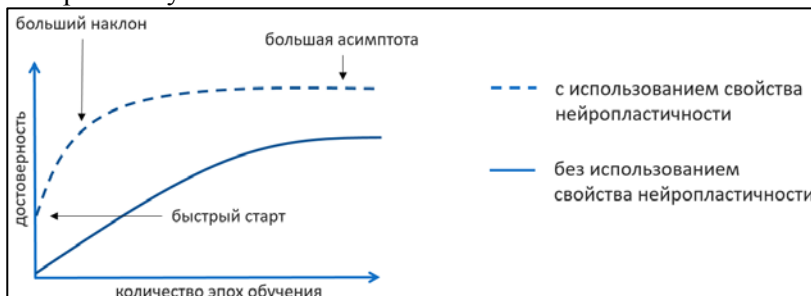


Рисунок 3. Визуализация целей свойства нейропластичности.

2. Нормализованная инициализация весов

В 2010 году в работе «Understanding The Difficulty Of Training Deep Feedforward Neural Networks» авторы провели исследование влияния методов начальной инициализации весов нейронных сетей и функций активации на обучение сетей [10]. В результате своей работы авторы предложили новый метод инициализации весов, который они назвали нормализованной инициализацией. В этом методе начальные веса определяются по следующей формуле:

$$W \sim U \left[-\frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}, \frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}} \right] \quad (1)$$

В этой формуле U – равномерное распределение на отрезке, n_j – количество нейронов в текущем слое сети, n_{j+1} – количество нейронов в следующем слое сети. Использование данного метода инициализации весов помогает снизить влияние проблемы исчезающего градиента, в результате чего сигнал ошибки распространяется лучше от выходного слоя к верхним слоям.

В 2015 метод был адаптирован для полулинейной функции активации ReLU, и веса определяются по новой формуле [11]:

$$W \sim U \left[-\frac{2}{n_j}, \frac{2}{n_j} \right] \quad (2)$$

Использование метода нормализованной инициализации весов позволяет качественно обучать глубокие нейронные без использования предобучения сети.

3. Обучение моделей нейронных сетей с использованием свойства нейропластичности

3.1. Описание данных для обучения

Для исследования свойства нейропластичности в свёрточных нейронных сетях в качестве данных для обучения была взята база изображений клеток крови [12]. Эта база разбита на 4 класса по подтипам клеток крови, которые находятся на изображении: эозинофилы, лимфоциты, моноциты, нейтрофилы.

Все изображения в базе являются маркированными, т.е. клетка, по которой должна производиться классификация выделена фиолетовым цветом.

База состоит из 12500 изображений клеток крови с соответствующими метками, относящими изображения к одному из классов. Каждое изображение имеет размер 320 на 240 пикселей и является цветным. Для исследовательской работы было решено использовать 3200 изображений для обучения, по 800 из каждого класса, 400 изображений для промежуточной проверки и 400 изображений для тестирования, по 100 изображений из каждого класса. Изображения будут сжаты до размеров 100 на 100 пикселей для сокращения времени обучения модели нейронной сети.

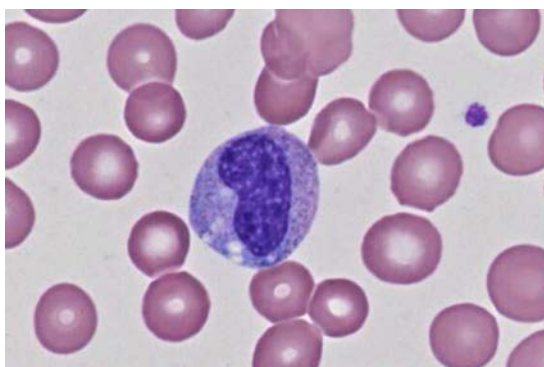


Рисунок 4. Маркированная клетка крови для классификации.

3.2. Описание эталонной модели нейронной сети

В качестве эталонной модели свёрточной нейронной сети было решено взять сеть VGG-16, обученную на наборе данных ImageNet [13]. Сеть VGG-16 содержит 16 слоёв и состоит из двух частей [14].

Первая часть выделяет характерные признаки в изображении. Она состоит из чередующихся каскадов свёрточных слоёв и слоёв подвыборки. Сеть состоит из двух каскадов свёртка-свёртка-подвыборка и трёх каскадов свёртка-свёртка-свёртка-подвыборка, которые идут друг за другом. Во всех свёрточных слоях размер ядра свёртки 3x3, а в качестве функции активации используется полулинейная функция ReLU. Слои подвыборки реализуют выбор максимального значения из окна изображения размером 2x2.

Вторая часть выполняет классификацию объекта на изображении по выделенным признакам на предыдущих слоях. Она состоит из трёх полносвязных слоёв: в первых двух слоях по 4096 нейронов, а в последнем 1000 нейронов.

Для решения новой задачи классификации подтипов клеток крови, к сети VGG-16 был добавлен полносвязный слой, состоящий из 256 нейронов с функцией активации ReLU, слой исключения с вероятностью отключения нейронов 0.5, полносвязный слой из 4-х нейронов (4 подтипа клеток крови) и функцией активацией SoftMax для классификации (рисунок 5).

3.3. Сравнение моделей, обученных с помощью свойства нейропластичности и случайной инициализации весов

Для исследования свойства нейропластичности свёрточных нейронных сетей было решено провести две серии экспериментов:

- обучить модель нейронной сети, в которой все слои инициализированы весами с использованием нормализованной инициализацией весов;
- обучить модель нейронной сети, в которой все слои инициализированы весами сети VGG-16, обученной на наборе данных ImageNet.

В качестве параметров обучения модели были выбраны следующие значения:

- количество эпох обучения 15;
- размер мини-выборки 20;
- в качестве функции потерь используется категориальная перекрёстная энтропия;
- в качестве метрики используется точность распознавания;
- в качестве оптимизатора используется метод адаптивной инерции (Adam).

Остальные параметры моделей для обучения соответствуют параметрам классической модели сети VGG-16, которую использовали для обучения на наборе данных ImageNet [14].

Для проверки хода обучения модели используется промежуточная проверочная выборка, которая подаётся на вход нейронной сети после окончания каждой эпохи обучения. Благодаря этому можно посмотреть, как менялась точность распознавания изображений после каждой эпохи. Под точностью понимают долю правильно распознанных данных подаваемых на вход сети.



Рисунок 5. Архитектура сверточной сети VGG-10 для исследования.



Рисунок 6. График изменения точности при обучении моделей.

На рисунке 6 изображен график изменения точности распознавания для двух серий эксперимента.

На тестовой выборке первая модель распознала данные с точностью в 0,75, а вторая модель распознала с точностью 0,985, что является очень хорошим результатом для распознавания изображений.

Как видно из результатов обучения, сеть, обученная с помощью свойства нейропластичности, показала лучший результат по сравнению с сетью, чьи веса были инициализированы случайным образом. Вторая модель уже на первых эпохах обучения имела высокую точность распознавания, а также она быстрее достигла своей верхней границы качества.

По результирующим данным можно сделать вывод, что первая модель за 15 эпох обучения не достигла своей верхней границы качества. Поэтому было решено обучить первую модель в течение большего количества эпох, чтобы лучше её исследовать.

Модель обучалась в течение 50 эпох и на тестовой выборке показала точность 0,9075. На рисунке 7 изображён график изменения точности распознавания модели во время обучения.



Рисунок 7. График изменения точности во время обучения первой модели.

Из результатов обучения первой модели можно сделать вывод, что даже при увеличении количества эпох обучения первой модели до 50, нейронная сеть не смогла достичь точности распознавания, которую достигла вторая модель за меньшее время обучения. Для достижения точности в 0,9 на промежуточной проверочной выборке первой модели потребовалось около 40 эпох, в то время как вторая модель достигла этого результата уже через 3 эпохи обучения.

3.4. Исследование свойства нейропластичности при различной инициализации слоёв нейронной сети

При использовании свойства нейропластичности для обучения модели нейронной сети не обязательно инициализировать все слои сети подготовленными весами. Можно загрузить подготовленные веса только в верхние слои сети, а глубокие слои инициализировать случайным образом.

Сеть VGG-16 состоит из 5 блоков свёрточных слоёв, которые разделены слоями подвыборки (рисунок 5). Для исследования нейропластичности было решено дополнительно обучить четыре модели сети, в дополнение к предыдущим экспериментам:

- в первый свёрточный блок загружены подготовленные веса, а веса остальных слоёв инициализированы с помощью нормализованной инициализации;
- в первый и второй свёрточный блок загружены подготовленные веса, а веса остальных слоёв инициализированы с помощью нормализованной инициализации;
- в первый, второй и третий свёрточный блок загружены подготовленные веса, а веса остальных слоёв инициализированы с помощью нормализованной инициализации;
- во все блоки, кроме пятого загружены подготовленные веса, а веса пятого блока инициализированы с помощью нормализованной инициализации.

Новые модели нейронных сетей были обучены в течение 15 эпох с теми же параметрами, что и две предыдущие модели. Для удобства обозначения новые модели будем называть третья, четвёртая, пятая и шестая. На рисунках 8, 9, 10 и 11 приведены графики изменения точности распознавания изображений для моделей во время обучения и их сравнение с первой и второй моделью сетей.

На тестовой выборке модели показали точность распознавания 0,8375, 0,85, 0,9175 и 0,9825 соответственно. Исходя из результатов, можно сделать вывод, что все модели показали результат обучения лучше, чем модель со случайной инициализацией весов во всех слоях сети. С увеличением количества слоёв, в которых веса инициализированы с помощью свойства нейропластичности, модель сети показывает большую точность распознавания изображений.

Таким образом, для решения поставленной задачи лучшей является модель нейронной сети, в которой всё свёрточные слои были инициализированы с помощью свойства

нейропластичности. Такая модель показала точность распознавания на тестовой выборке 0,985, что является лучшим результатом среди всех остальных моделей сетей.

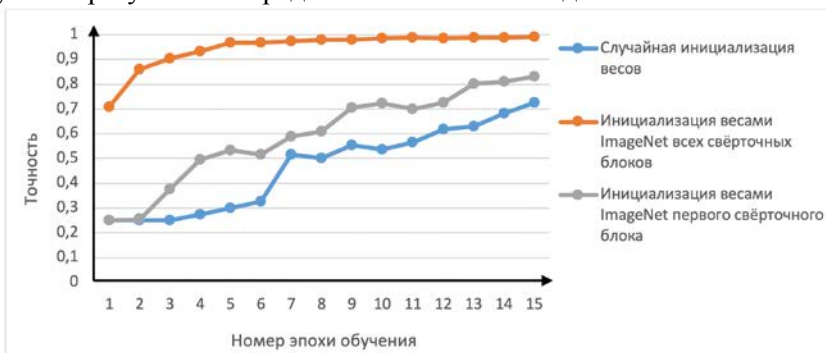


Рисунок 8. График изменения точности распознавания изображений во время обучения третьей модели.

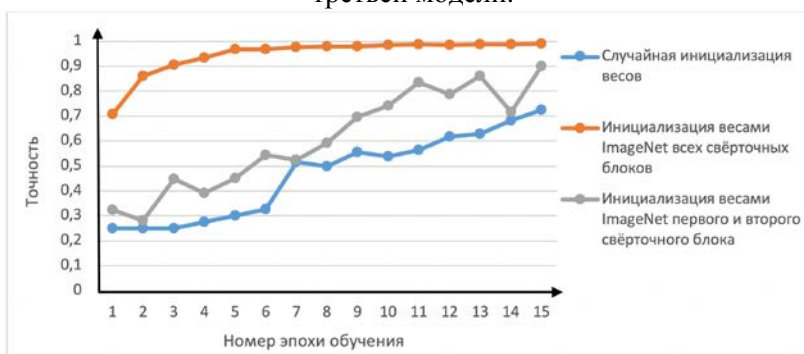


Рисунок 9. График изменения точности распознавания изображений во время обучения четвертой модели.

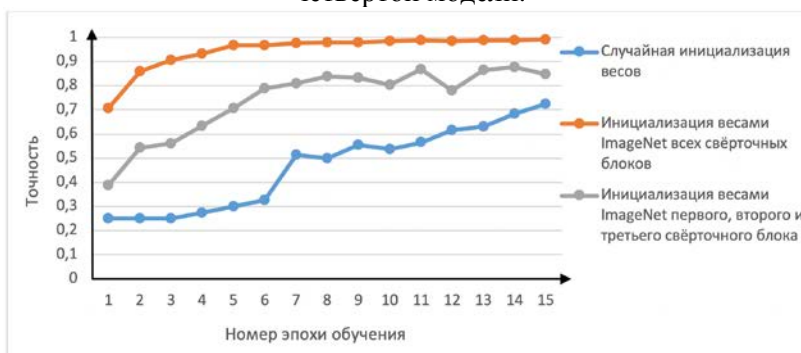


Рисунок 10. График изменения точности распознавания изображений во время обучения пятой модели.

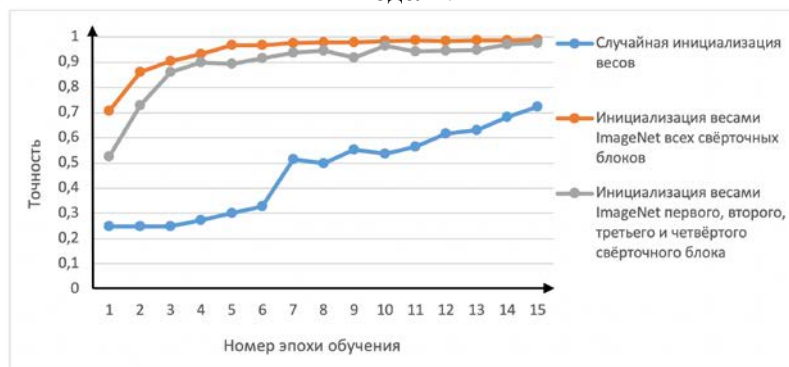


Рисунок 11. График изменения точности распознавания изображений во время обучения шестой модели.

4. Заключение

В статье было проведено исследование свойства нейропластичности в свёрточных нейронных сетях на примере решения задачи классификации подтипов клеток крови.

При проведении исследования обучения разных моделей была найдена конфигурация инициализации весов, которая дала точность распознавания изображений 0,985. В этой модели во всех свёрточных слоях веса были инициализированы с использованием свойства нейропластичности. Самый худший результат показала модель, в которой все веса в свёрточных слоях были инициализированы случайным образом.

Дополнительные исследования, которые заключались в использовании комбинации способов инициализации весов для обучения, показали, что с увеличением количества слоёв, в которых веса инициализированы с помощью свойства нейропластичности, модели дают лучшую точность распознавания изображений.

5. Литература

- [1] Torrey, L. Transfer Learning / L. Torrey, J. Shavlik // Handbook of Research on Machine Learning Applications. – IGI Global, 2009. – P. 245-267.
- [2] Medical Definition of Neuroplasticity // MedicineNet [Electronic resource]. – Access mode: <https://www.medicinenet.com/script/main/art.asp?articlekey=40362> (23.01.2018).
- [3] Doidge, N. The Brain That Changes Itself // Penguin Books, 2007. – 224 p.
- [4] Faiz, M. Proliferation Dynamics Of Germinative Zone Cells In The Intact and excitotoxic ally lesioned postnatal rat brain / M. Faiz, L. Acarin, B. Casteallano, B. Gonzalez // BMC Neurosci, 2005. – P. 687-702.
- [5] LeCun, Y. Backpropagation Applied to Handwritten Zip Code Recognition / Y. LeCun, B. Boser, D. Henderson // Neural Computation, 1989. – P. 541-551.
- [6] Zeiler, M. Visualizing and Understanding Convolution Networks / M. Zeiler, R. Fergus // Computing Research Repository, 2013. – 16 p.
- [7] Сойфер, В.А. Методы компьютерной обработки изображений // Физматлит, 2003. – 459 с.
- [8] Галушкин, А.И. Синтез многослойных систем распознавания образов // Энергия, 1974. – 368 с.
- [9] Hochreiter, S. Gradient Flow In Recurrent Nets: The Difficulty Of Learning Long-Term Dependencies / S. Hochreiter, Y. Bengio // A Field Guide to Dynamical Recurrent Neural Networks, 2001. – 464 p.
- [10] Glorot, X. Understanding The Difficulty Of Training Deep Feedforward Neural Networks / X. Glorot, Y. Bengio // Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 2010. – P. 8.
- [11] He, K. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification / K. He, X. Zhang // Proceedings of the IEEE International Conference on Computer Vision, 2015. – P. 8.
- [12] База данных клеток крови // Kaggle [Электронный ресурс]. – Режим доступа: <https://www.kaggle.com/paultimothymooney/blood-cells> (27.09.2018).
- [13] About ImageNet // ImageNet [Электронный ресурс]. – Режим доступа: <http://imagenet.org/about-overview> (2.02.2018).
- [14] Simonyan, K. Very Deep Convolutional Networks for Large-Scale Image Recognition / K. Simonyan, A. Zisserman // Computer Vision and Pattern Recognition, 2014. – 14 p.

Investigation of the effectiveness of using the property of neuroplasticity in convolutional networks

G.A. Algashev¹, O.P. Soldatova¹

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. When training neural networks, one can often encounter the following problems: a small training sample, a long time of training the neural network, the problem of a vanishing gradient. These problems have long been an obstacle to learning really deep neural networks. To solve the problems posed, the transfer learning method based on the neuroplasticity property of neural networks has been developed. The idea of the neuroplasticity property in convolutional networks is to apply a new class of images to a well-trained network, which the network did not learn before, and continue to train the network with a new set of data. In the article, the study of the properties of neuroplasticity occurs on the example of solving the problem of classification of marked blood cells. For the study, several models of neural networks are trained using the method of normalized initialization of weights, the method of transfer learning and the method of mixed initialization of weights. With an increase in the number of layers in which weights are initialized using the transfer learning method, the reliability of the models increases. The best result was shown by a model in which all layers were initialized using the neuroplasticity property.