$$E = \frac{1}{2}\sum_{j=1}^{n}(\bar{x}_i - x_i)^2 . \qquad (40)$$

Тогда в соответствии с методом градиентного спуска

$$\frac{\partial E}{\partial w'_{ji}} = \frac{\partial E}{\partial \bar{x}_i}\frac{\partial \bar{x}_i}{\partial w'_{ji}} = = -(x_i - \bar{x}_i)y_j = -y_j(x_i - \sum_{j=1}^{p} w'_{ji}y_j) . \qquad (41)$$

Отсюда получается Ойя правило для обучения рециркуляционной сети

$$w'_{ji}(t+1) = w'_{ji}(t) + \alpha y_j(x_i - \sum_{j=1}^{p} w'_{ji}y_j) . \qquad (42)$$

В соответствии с методом главных компонент [9] матрица весовых коэффициентов W прямого слоя

$$W = (W')^T \qquad (43)$$

Тогда правило изменения весовых коэффициентов прямого слоя можно представить как

$$w_{ij}(t+1) = w_{ij}(t) + \alpha y_j\left(x_i - \sum_{j=1}^{p} w_{ij}y_j\right) . \qquad (44)$$

Теоретически доказано [16], что Ойя-правило эквивалентно методу главных компонент, однако его применение приводит к слишком длительной процедуре обучения.

Заключение

В работе рассмотрены основные парадигмы обучения нейронных сетей: обучение с учителем, подкрепляющее обучение и обучение без учителя. Проведен анализ алгоритмов обратного распространения ошибки, АНС и Q-обучения, правила Хебба, конкурентного обучения и правила Ойя.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. MCculloh W., Pitts W. A logical calculus of the ideas immanent in nervous activity // Bulletin of Mathematical Biophysics. - 1943. - N5. - P. 115 - 133.
2. Widrow B., Hoff M. Adaptive switching circuits// In 1960 IRE WESCON Convention Record. – DUNNO, 1960. - P. 96 - 104.
3. Minsky M., Papert S. Perceptrons: An introduction to Computational Geometry. - The MIT Press, 1969. – 153 p.
4. Rumelhart D., Hinton G., Williams R. Learning internal representations by error propagation // Parallel Distributed Processing. – 1986. – Vol. 1. – P. 318 - 362
5. Hopfield J. Neural Networks and physical systems with emergent collective computational abilities // Proceedings of the National Academy of Sciences USA. - 1982. - P. 2554 - 2558.
6. Kohonen T. Self-organised formation of topologically correct feature maps // Biological Cybernetics. - 1982. – N. 43. - P. 59 - 69.
7. Головко В.А. Нейроинтеллект: теория и применение. Книга 1. Организация и обучение нейронных сетей с прямыми и обратными связями. – Брест: Изд. БПИ, 1999. – 264 с.
8. Головко В.А. Нейроинтеллект: теория и применение. Книга 2. Самоорганизация, отказоустойчивость и применение нейронных сетей. – Брест: Изд. БПИ, 1999. – 228 с.
9. Головко В.А. Нейронные сети: обучение, организация и применение. Кн. 4: Учеб. пособие для вузов / Общая ред. А.И. Галушкина. – М.: ИПРЖР, 2001. – 256 с.
10. Schiffman W., Joost M., Werner R. Comparison of optimized backpropagation algorithms // Proc. of the European Symposium on Artificial Neural Networks, ESANN'93. - Brussels, 1992. – P. 97 - 104.
11. Silva F., Almeida L. Speeding up Back-Propagation // Advanced Neural Computers. – 1990. – P. 151 - 158.
12. V. Golovko, N. Maniakov, L. Makhnist. Multilayer neural networks training methodic // Proceedings of the Second IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. (IDAACS'2003), Lviv, Ukraine, September 8-10, 2003. – P. 185-190.
13. Sutton R., Barto A. Reinforcement learning. – Cambridge MA: MIT Press, 1995.
14. Kohonen T. Self-organizing and Associative Memory. - Berlin: Springer-Verlag, 1984. – 135 p.
15. Taylor J. Towards the networks of the brain: from brain imaging to consciousness // Neural Networks. – 1999 – Vol. 12, N. 7 - 8. – P. 943 - 959.
16. Oja E. A simplified neuron model as a principal component analyzer // Journal of mathematical biology. – 1982. – N. 15. – P. 267 - 273.

УДК 681.324:519.711.7

Гладкий И.И., Маньяков Н.В., Махнист Л.П.

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ АЛГОРИТМОВ ОБУЧЕНИЯ ГЕТЕРОГЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДВУХСЛОЙНОЙ АРХИТЕКТУРЫ

1. ВВЕДЕНИЕ

Двухслойные нейронные гетерогенные сети прямого распространения являются наиболее используемым классом нейронных сетей. Более 80% прикладных задач, решаемых в нейросетевом базисе, используют данный тип архитектуры [1]. Широкое применение таких сетей основано на способности данного класса сетей быть универсальным аппроксиматором [2], что позволяет любую задачу приближения решать с их использованием. Их применение встречается в задачах прогнозирования, классификации, кластеризации, управления и многих других. В связи с этим возникает вопрос о создании эффективной методики их обучения.

2. ЗАДАЧА ОБУЧЕНИЯ ДВУХСЛОЙНОЙ СЕТИ И ЕЕ МАТРИЧНАЯ АЛГОРИТМИЗАЦИЯ

Рассмотрим двухслойную гетерогенную нейронную сеть, состоящую из m_0 нейронных элементов распределительного слоя, m_1 нейронов скрытого слоя и m_2 – выходного слоя (рис.1).

Гладкий Иван Иванович, *ст. преподаватель каф. высшей математики Брестского государственного технического университета.*

Маньяков Николай Владимирович, *ст. преподаватель каф. высшей математики Брестского государственного технического университета.*

Махнист Леонид Петрович, *к.т.н., доцент каф. высшей математики Брестского государственного технического университета.*

Беларусь, БГТУ, 224017, г. Брест, ул. Московская, 267.

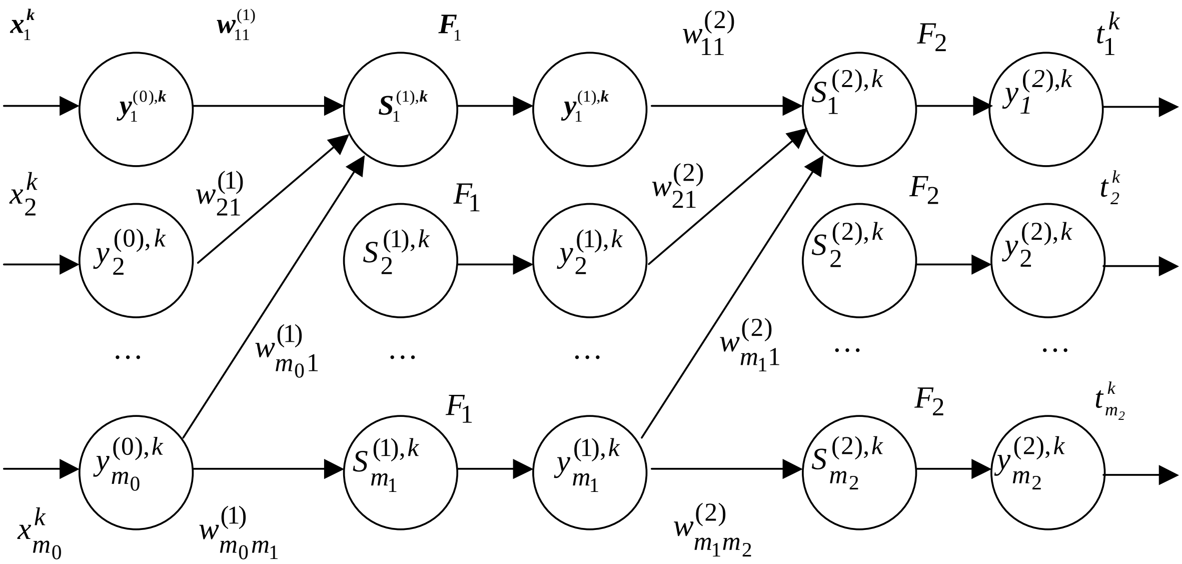

Рис. 1. Схема функционирования нейронной сети.

Для данной сети каждый нейрон распределительного слоя имеет синаптические связи $w^{(1)}_{i_0 i_1}$ $\left(i_0=\overline{1,m_0},\ i_1=\overline{1,m_1}\right)$ со всеми нейронами скрытого слоя, а каждый нейрон скрытого слоя синаптические связи $w^{(2)}_{i_1 i_2}$ $\left(i_1=\overline{1,m_1},\ i_2=\overline{1,m_2}\right)$ со всеми нейронами выходного слоя. В качестве нейронов скрытого слоя используются элементы с функцией активации F_1, в качестве нейронов выходного слоя – с функцией активации F_2. На вход сети подаются входные образы – векторы $\overline{x^k}=\left(x_1^k,...,x_{m_0}^k\right)^T$, $\left(k=\overline{1,L}\right)$. Выходами распределительного слоя являются значения $y_{i_0}^{(0),k}=x_{i_0}^k$. При этом формируется вектор $Y^{(0),k}=\left(y_1^{(0),k}\quad y_2^{(0),k}\quad \dots\quad y_{m_0}^{(0),k}\quad -1\right)^T$.

Выходное значение i_1-ого нейрона скрытого слоя сети для k-ого образа определяется соотношением

$$y_{i_1}^{(1),k}=F_1\left(S_{i_1}^{(1),k}\right),$$

где

$$S_{i_1}^{(1),k}=\sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(0),k}-T_{i_1}^{(1)},\quad i_1=\overline{1,m_1},\quad k=\overline{1,L}.$$

При этом формируется вектор $Y^{(1),k}=\left(y_1^{(1),k}\quad y_2^{(1),k}\quad \dots\quad y_{m_1}^{(1),k}\quad -1\right)^T$.

Выходное значение i_2-ого нейрона выходного слоя сети для k-ого образа определяется соотношением

$$y_{i_2}^{(2),k}=F_2\left(S_{i_2}^{(2),k}\right),$$

где

$$S_{i_2}^{(2),k}=\sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} y_{i_1}^{(1),k}-T_{i_2}^{(2)},\quad i_2=\overline{1,m_2},\quad k=\overline{1,L}.$$

Задача обучения нейронной сети с фиксированными функциями активации [3] состоит в нахождении весовых коэффициентов $w_{i_0 i_1}^{(1)}\left(i_0=\overline{1,m_0},\ i_1=\overline{1,m_1}\right)$, $w_{i_1 i_2}^{(2)}\left(i_1=\overline{1,m_1},\ i_2=\overline{1,m_2}\right)$ и порогов нейронных элементов $T_{i_1}^{(1)}\left(i_1=\overline{1,m_1}\right)$, $T_{i_2}^{(2)}\left(i_2=\overline{1,m_2}\right)$, которые минимизируют некоторую ошибку сети E_S, как отклонение выходных значений $y_{i_2}^{(2),k}$ от эталонных значений $t_{i_2}^k$ – i_2-ого нейрона сети для k-ого образа. В качестве ошибки будем рассматривать усредненное по количеству образов «квадратичное отклонение»

$$E_S=\frac{1}{2L}\sum_{k=1}^{L}\sum_{i_2=1}^{m_2}\left(y_{i_2}^{(2),k}-t_{i_2}^k\right)^2.$$

Матрицы $W^{(1)}=\begin{pmatrix} w_{11}^{(1)} & w_{21}^{(1)} & \dots & w_{m_0 1}^{(1)} \\ w_{12}^{(1)} & w_{22}^{(1)} & \dots & w_{m_0 2}^{(1)} \\ \dots & \dots & \dots & \dots \\ w_{1m_1}^{(1)} & w_{2m_1}^{(1)} & \dots & w_{m_0 m_1}^{(1)} \end{pmatrix}_{m_1\times m_0}$ и

$W^{(2)}=\begin{pmatrix} w_{11}^{(2)} & w_{21}^{(2)} & \dots & w_{m_1 1}^{(2)} \\ w_{12}^{(2)} & w_{22}^{(2)} & \dots & w_{m_1 2}^{(2)} \\ \dots & \dots & \dots & \dots \\ w_{1m_2}^{(2)} & w_{2m_2}^{(2)} & \dots & w_{m_1 m_2}^{(2)} \end{pmatrix}_{m_2\times m_1}$ и векторы

$\overline{T^{(1)}}=\left(T_1^{(1)},\ T_2^{(1)},...,\ T_{m_1}^{(1)}\right)^T$, $\overline{T^{(2)}}=\left(T_1^{(2)},\ T_2^{(2)},...,\ T_{m_2}^{(2)}\right)^T$

будем называть приближенным решением или просто решением (по методу наименьших квадратов) системы

$$F_2\left(\sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)}\cdot F_1\left(\sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(0),k}-T_{i_1}^{(1)}\right)-T_{i_2}^{(2)}\right)=t_{i_2}^k,$$

$$i_2=\overline{1,m_2},\quad k=\overline{1,L}$$

если «квадратичное отклонение»

$$E_S=\frac{1}{2L}\sum_{k=1}^{L}\sum_{i_2=1}^{m_2}\left(F_2\left(\left(\sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)}\cdot F_1\left(\sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(1),k}-T_{i_1}^{(1)}\right)-T_{i_2}^{(2)}\right)-t_{i_2}^k\right)^2=\frac{1}{L}\sum_{k=1}^{L}E_s^{(k)}$$

достигает своего наименьшего значения, где

$$E_s^{(k)} = \frac{1}{L}\cdot\sum_{i_2=1}^{m_2}\left(F_2\left(\sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)}\cdot F_1\left(\sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(1),k} - T_{i_1}^{(1)}\right) - T_{i_2}^{(2)}\right) - t_{i_2}^{k}\right)^2$$

Рассмотрим схему алгоритмов обучения двухслойной сети с использованием градиентных методов.

В соответствии с градиентными методами изменение весов и порогов сети производится по формулам:

$$w_{j_0 j_1}^{(1)}(t+1) = w_{j_0 j_1}^{(1)}(t) - \alpha^{(1)}\frac{\partial E_s(t)}{\partial w_{j_0 j_1}^{(1)}},$$

$$j_0 = \overline{1,m_0},\quad j_1 = \overline{1,m_1},$$

$$T_{j_1}^{(1)}(t+1) = T_{j_1}^{(1)}(t) - \alpha^{(1)}\frac{\partial E_s(t)}{\partial T_{j_1}^{(1)}},\qquad j_1 = \overline{1,m_1},$$

$$w_{j_1 j_2}^{(2)}(t+1) = w_{j_1 j_2}^{(2)}(t) - \alpha^{(2)}\frac{\partial E_s(t)}{\partial w_{j_1 j_2}^{(2)}},$$

$$j_1 = \overline{1,m_1},\quad j_2 = \overline{1,m_2},$$

$$T_{j_2}^{(2)}(t+1) = T_{j_2}^{(2)}(t) - \alpha^{(2)}\frac{\partial E_s(t)}{\partial T_{j_2}^{(2)}},\qquad j_2 = \overline{1,m_2},$$

где $\dfrac{\partial E_s}{\partial z} = \dfrac{\partial\left(\frac{1}{L}\sum_{k=1}^{L} E_s^{(k)}\right)}{\partial z} = \dfrac{1}{L}\sum_{k=1}^{L}\dfrac{\partial E_s^{(k)}}{\partial z}$ и $\alpha^{(1)}$, $\alpha^{(2)}$ – параметры обучения, выбираемые постоянными или зависящими от t (адаптивными), для достижения наименьшего значения ошибки.

В соответствии с принципом матричного метода обучения сети [4], соотношения для изменения синаптических связей запишутся в виде:

$$w_{j_1 j_2}^{(2)}(t+1) = w_{j_1 j_2}^{(2)}(t) - \alpha^{(2)}\cdot\frac{1}{L}\cdot G_{j_1 j_2}^{(2)},\quad j_1 = \overline{1,m_1},\quad j_2 = \overline{1,m_2},$$

$$T_{j_2}^{(2)}(t+1) = T_{j_2}^{(2)}(t) - \alpha^{(2)}\cdot\frac{1}{L}\cdot G_{(m_1+1) j_2}^{(2)},\qquad j_2 = \overline{1,m_2},$$

$$w_{j_0 j_1}^{(1)}(t+1) = w_{j_0 j_1}^{(1)}(t) - \alpha^{(1)}\cdot\frac{1}{L}\cdot G_{j_0 j_1}^{(1)},$$

$$T_{j_1}^{(1)}(t+1) = T_{j_1}^{(1)}(t) - \alpha^{(1)}\cdot\frac{1}{L}\cdot G_{(m_0+1) j_1}^{(1)},$$

где $G_{j_1 j_2}^{(2)} = \sum_{k=1}^{L} C^{(2),k}\cdot K_{j_1 j_2}^{(2),k}$, $G_{j_0 j_1}^{(1)} = \sum_{k=1}^{L} C^{(1),k}\cdot K_{j_0 j_1}^{(1),k}$ и

$$C^{(2),k} = \varepsilon_2^k\cdot MF_2',\qquad K_{j_1 j_2}^{(2),k} = M_{j_2 j_1}^{(2)}\cdot Y^{(1),k},$$

$$C^{(1),k} = C^{(2),k}\cdot W^{(2)}\cdot MF_1',\quad K_{j_0 j_1}^{(1),k} = M_{j_1 j_0}^{(1)}\cdot Y^{(0),k}.$$

Здесь и далее принимаем следующие обозначения

$$\varepsilon_2^k = \left(\left(y_1^{(2),k} - t_1^k\right)\ \left(y_2^{(2),k} - t_2^k\right)\ \ldots\ \left(y_{m_2}^{(2),k} - t_{m_2}^k\right)\right),$$

$$MF_n' = \begin{pmatrix} F_n'\left(S_1^{(n),k}\right) & 0 & \ldots & 0 \\ 0 & F_n'\left(S_2^{(n),k}\right) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & F_n'\left(S_{m_n}^{(n),k}\right) \end{pmatrix},$$

$n = 1,2$ – матрица размерности $m_n\times m_n$, а матрица $M_{j_n j_{n-1}}^{(n)}$ размерности $m_n\times(m_{n-1}+1)$ состоит из числа 1 на позиции $j_n j_{n-1}$ и нулей в качестве остальных элементов матрицы.

3. МЕТОДИКИ ВЫБОРА АДАПТИВНОГО ШАГА ОБУЧЕНИЯ

Для двухслойной нелинейной гетерогенной нейронной сети можно предложить следующие различные методы обучения с использованием адаптивного шага на основе метода наискорейшего спуска:

1. Послойное обучение. Данный метод заключается в том, что сначала, после подачи всех элементов обучающего множества, изменяются синаптические связи последнего (или первого) слоя, а после того как на измененную сеть были опять поданы все элементы обучающего множества, изменябтся синаптические связи оставшегося слоя. Затем осуществляем переход к началу алгоритма. Процедура обучения останавливается после того, как ошибка сети не превышает заданную после двух последовательных итераций. Шаги обучения $\alpha^{(1)}$ и $\alpha^{(2)}$ выбираются для наилучшей минимизации ошибки сети при изменении синаптических связей соответствующих слоев.
2. Двухпараметрическое обучение. Данный метод заключается в том, что после подачи всех элементов обучающего подмножество отыскиваются такие адаптивные шаги обучения $\alpha^{(1)}$ и $\alpha^{(2)}$, с использованием которых минимизируется функцию ошибки обучения сети.
3. Обобщенный метод наискорейшего спуска. При данном методе, после подачи всех элементов обучающего подмножества ищется такой адаптивный шаг обучения $\alpha = \alpha^{(1)} = \alpha^{(2)}$, который минимизирует ошибку сети в направлении антиградиента.

Для каждого из приведённых методов адаптивный шаг вычисляется в соответствии, с приведенными ниже теоремами. Для их доказательства используются следующие соотношения [5]:

$$\left(S^{(2)}\right)_{l_1 l_2}^{j_1 j_2} = \sum_{k=1}^{L}\left(\left(K_{l_1 l_2}^{(2),k}\right)^T\cdot\left(\left(MF_2'\right)^2 + DE^k\cdot MF_2''\right)\cdot K_{j_1 j_2}^{(2),k}\right),$$

$$\left(S^{(1)}\right)_{l_0 l_1}^{j_0 j_1} = \sum_{k=1}^{L}\left(\left(K_{l_0 l_1}^{(1),k}\right)^T\cdot\left(\left(MW^{(1)}\right)^T\cdot\left[\left(MF_2'\right)^2 + DE^k\cdot MF_2''\right]\cdot MW^{(1)} + \right.\right.$$

$$\left.\left. + DE^k\cdot\left(MF_2'\cdot W^{(2)}\right)\cdot MF_1''\right)\cdot K_{j_0 j_1}^{(1),k}\right),$$

$$\left(S^{(1,2)}\right)_{l_1 l_2}^{j_0 j_1} = \sum_{k=1}^{L}\left(\left(K_{l_1 l_2}^{(2),k}\right)^T\cdot\left(\left(MF_2'\right)^2 + DE^k\cdot MF_2''\right)\cdot MW^{(1)}\cdot K_{j_0 j_1}^{(1),k}\right)$$

.

Теорема 1. При послойном обучении двухслойной гетерогенной нейронной сети адаптивный шаг для каждого слоя определяется в соответствии с формулами:

$$\alpha_m^{(2)} = \frac{\sum_{j_1=1}^{m_1+1}\sum_{j_2=1}^{m_2}\left(G_{j_1 j_2}^{(2)}\right)^2}{\sum_{j_1,l_1=1}^{m_1+1}\sum_{j_2,l_2=1}^{m_2} G_{l_1 l_2}^{(2)}\cdot\left(S^{(2)}\right)_{l_1 l_2}^{j_1 j_2} G_{j_1 j_2}^{(2)}},$$

$$\alpha_m^{(1)} = \frac{\sum_{j_0=1}^{m_0+1}\sum_{j_1=1}^{m_1}\left(G_{j_0 j_1}^{(1)}\right)^2}{\sum_{j_0,l_0=1}^{m_0+1}\sum_{j_1,l_1=1}^{m_1} G_{l_0 l_1}^{(1)}\cdot\left(S^{(1)}\right)_{l_0 l_1}^{j_0 j_1} G_{j_0 j_1}^{(1)}}.$$

При этом модификация синаптических связей происходит по формулам:

$$w_{j_1 j_2}^{(2)}(t+1) = w_{j_1 j_2}^{(2)}(t) - \alpha_m^{(2)} \cdot G_{j_1 j_2}^{(2)}, \quad j_1 = \overline{1, m_1}, \quad j_2 = \overline{1, m_2}$$

$$T_{j_2}^{(2)}(t+1) = T_{j_2}^{(2)}(t) - \alpha_m^{(2)} \cdot G_{(m_1+1) j_2}^{(2)}, \qquad j_2 = \overline{1, m_2}$$

$$w_{j_0 j_1}^{(1)}(t+1) = w_{j_0 j_1}^{(1)}(t) - \alpha_m^{(1)} \cdot G_{j_0 j_1}^{(1)}, \quad j_0 = \overline{1, m_0}, \quad j_1 = \overline{1, m_1}$$

$$T_{j_1}^{(1)}(t+1) = T_{j_1}^{(1)}(t) - \alpha_m^{(1)} \cdot G_{(m_0+1) j_1}^{(1)}, \qquad j_1 = \overline{1, m_1},$$

Доказательство.

При изменение синаптических связей второго слоя, ошибка сети меняется в соответствии с формулой

$$E_s(t+1) = \frac{1}{L} \cdot \sum_{k=1}^{L} E_s^{(k)}(t+1) = \frac{1}{L} \sum_{k=1}^{L} E_s^{(k)}(t) +$$

$$+\frac{1}{L} \cdot \left(\sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \left(\sum_{k=1}^{L} \frac{\partial E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}} \right) \cdot \left(w_{j_1 j_2}^{(2)}(t+1) - w_{j_1 j_2}^{(2)}(t) \right) + \right.$$

$$\left. + \sum_{j_2=1}^{m_2} \left(\sum_{k=1}^{L} \frac{\partial E_s^{(k)}}{\partial T_{j_2}^{(2)}} \right) \cdot \left(T_{j_2}^{(2)}(t+1) - T_{j_2}^{(2)}(t) \right) \right) +$$

$$+\frac{1}{2L} \cdot \sum_{k=1}^{L} \left(\sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \sum_{l_1=1}^{m_1} \sum_{l_2=1}^{m_2} \frac{\partial^2 E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)} \partial w_{l_1 l_2}^{(2)}} \cdot \left(w_{j_1 j_2}^{(2)}(t+1) - \right. \right.$$

$$\left. - w_{j_1 j_2}^{(2)}(t) \right) \cdot \left(w_{l_1 l_2}^{(2)}(t+1) - w_{l_1 l_2}^{(2)}(t) \right) +$$

$$+ \sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \sum_{l_2=1}^{m_2} \frac{\partial^2 E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)} \partial T_{l_2}^{(2)}} \cdot \left(w_{j_1 j_2}^{(2)}(t+1) - \right.$$

$$\left. - w_{j_1 j_2}^{(2)}(t) \right) \cdot \left(T_{l_2}^{(2)}(t+1) - T_{l_2}^{(2)}(t) \right) +$$

$$+ \sum_{l_1=1}^{m_1} \sum_{l_2=1}^{m_2} \sum_{j_2=1}^{m_2} \frac{\partial^2 E_s^{(k)}}{\partial T_{j_2}^{(2)} \partial w_{l_1 l_2}^{(2)}} \cdot \left(T_{j_2}^{(2)}(t+1) - \right.$$

$$\left. - T_{j_2}^{(2)}(t) \right) \cdot \left(w_{l_1 l_2}^{(2)}(t+1) - w_{l_1 l_2}^{(2)}(t) \right) +$$

$$+ \sum_{l_2=1}^{m_2} \sum_{j_2=1}^{m_2} \frac{\partial^2 E_s^{(k)}}{\partial T_{l_2}^{(2)} \partial T_{j_2}^{(2)}} \cdot \left(T_{l_2}^{(2)}(t+1) - \right.$$

$$\left. \left. - T_{l_2}^{(2)}(t) \right) \cdot \left(T_{j_2}^{(2)}(t+1) - T_{j_2}^{(2)}(t) \right) \right) =$$

$$= E_s(t) - \alpha^{(2)} \cdot \frac{1}{L^2} \cdot \sum_{j_1=1}^{m_1+1} \sum_{j_2=1}^{m_2} \left(G_{j_1 j_2}^{(2)} \right)^2 +$$

$$+ \left(\alpha^{(2)} \right)^2 \cdot \frac{1}{2L^3} \cdot \left(\sum_{j_1, l_1=1}^{m_1+1} \sum_{j_2, l_2=1}^{m_2} G_{l_1 l_2}^{(2)} \cdot \sum_{k=1}^{L} \left(\left(K_{l_1 l_2}^{(2),k} \right)^T \cdot \left(\left(MF_2' \right)^2 + \right. \right. \right.$$

$$\left. \left. \left. + DE^k \cdot MF_2'' \right) \cdot K_{j_1 j_2}^{(2),k} \right) \cdot G_{j_1 j_2}^{(2)} \right)$$

Найдем, при каком значении $\alpha^{(2)}$ полученное значение ошибки сети будет минимальным. Для этого решим уравнение

$$\frac{\partial E_s(t+1)}{\partial \alpha^{(2)}} = -\frac{1}{L^2} \cdot \sum_{j_1=1}^{m_1+1} \sum_{j_2=1}^{m_2} \left(G_{j_1 j_2}^{(2)} \right)^2 +$$

$$+ \alpha^{(2)} \cdot \frac{1}{L^3} \cdot \left(\sum_{j_1, l_1=1}^{m_1+1} \sum_{j_2, l_2=1}^{m_2} G_{l_1 l_2}^{(2)} \cdot \left(S^{(2)} \right)_{l_1 l_2}^{j_1 j_2} G_{j_1 j_2}^{(2)} \right) = 0.$$

Полученное значение $\alpha^{(2)}$ равно

$$\alpha^{(2)} = L \cdot \frac{\sum_{j_1=1}^{m_1+1} \sum_{j_2=1}^{m_2} \left(G_{j_1 j_2}^{(2)} \right)^2}{\sum_{j_1, l_1=1}^{m_1+1} \sum_{j_2, l_2=1}^{m_2} G_{l_1 l_2}^{(2)} \cdot \left(S^{(2)} \right)_{l_1 l_2}^{j_1 j_2} G_{j_1 j_2}^{(2)}}.$$

Таким образом

$$\alpha_m^{(2)} = \alpha^{(2)} \cdot \frac{1}{L} = \frac{\sum_{j_1=1}^{m_1+1} \sum_{j_2=1}^{m_2} \left(G_{j_1 j_2}^{(2)} \right)^2}{\sum_{j_1, l_1=1}^{m_1+1} \sum_{j_2, l_2=1}^{m_2} G_{l_1 l_2}^{(2)} \cdot \left(S^{(2)} \right)_{l_1 l_2}^{j_1 j_2} G_{j_1 j_2}^{(2)}}.$$

Модифицируя синаптические связи только первого слоя, ошибка сети изменяется по формуле

$$E_s(t+1) = \frac{1}{L} \cdot \sum_{k=1}^{L} E_s^{(k)}(t+1) = \frac{1}{L} \sum_{k=1}^{L} E_s^{(k)}(t) +$$

$$+\frac{1}{L} \cdot \left(\sum_{j_0=1}^{m_0} \sum_{j_1=1}^{m_1} \left(\sum_{k=1}^{L} \frac{\partial E_s^{(k)}}{\partial w_{j_0 j_1}^{(1)}} \right) \cdot \left(w_{j_0 j_1}^{(1)}(t+1) - w_{j_0 j_1}^{(1)}(t) \right) + \right.$$

$$\left. + \sum_{j_1=1}^{m_1} \left(\sum_{k=1}^{L} \frac{\partial E_s^{(k)}}{\partial T_{j_1}^{(1)}} \right) \cdot \left(T_{j_1}^{(1)}(t+1) - T_{j_1}^{(1)}(t) \right) \right) +$$

$$+\frac{1}{2L} \cdot \sum_{k=1}^{L} \left(\sum_{j_0=1}^{m_0} \sum_{j_1=1}^{m_1} \sum_{l_0=1}^{m_0} \sum_{l_1=1}^{m_1} \frac{\partial^2 E_s^{(k)}}{\partial w_{j_0 j_1}^{(1)} \partial w_{l_0 l_1}^{(1)}} \cdot \left(w_{j_0 j_1}^{(1)}(t+1) - \right. \right.$$

$$\left. - w_{j_0 j_1}^{(1)}(t) \right) \cdot \left(w_{l_0 l_1}^{(1)}(t+1) - w_{l_0 l_1}^{(1)}(t) \right) +$$

$$+ \sum_{j_0=1}^{m_0} \sum_{j_1=1}^{m_1} \sum_{l_1=1}^{m_1} \frac{\partial^2 E_s^{(k)}}{\partial w_{j_0 j_1}^{(1)} \partial T_{l_1}^{(1)}} \cdot \left(w_{j_0 j_1}^{(1)}(t+1) - \right.$$

$$\left. - w_{j_0 j_1}^{(1)}(t) \right) \cdot \left(T_{l_1}^{(1)}(t+1) - T_{l_1}^{(1)}(t) \right) +$$

$$+ \sum_{l_0=1}^{m_0} \sum_{l_1=1}^{m_1} \sum_{j_1=1}^{m_1} \frac{\partial^2 E_s^{(k)}}{\partial T_{j_1}^{(1)} \partial w_{l_0 l_1}^{(1)}} \cdot \left(T_{j_1}^{(1)}(t+1) - T_{j_1}^{(1)}(t) \right) \times$$

$$\times \left(w_{l_0 l_1}^{(1)}(t+1) - w_{l_0 l_1}^{(1)}(t) \right) +$$

$$+ \sum_{l_1=1}^{m_1} \sum_{j_1=1}^{m_1} \frac{\partial^2 E_s^{(k)}}{\partial T_{l_1}^{(1)} \partial T_{j_1}^{(1)}} \cdot \left(T_{l_1}^{(1)}(t+1) - \right.$$

$$\left. \left. - T_{l_1}^{(1)}(t) \right) \cdot \left(T_{j_1}^{(1)}(t+1) - T_{j_1}^{(1)}(t) \right) \right) =$$

$$= E_s(t) - \alpha^{(1)} \cdot \frac{1}{L^2} \cdot \sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \left(G_{j_0 j_1}^{(1)} \right)^2 +$$

$$+ \left(\alpha^{(1)} \right)^2 \cdot \frac{1}{2L^3} \cdot \left(\sum_{j_0, l_0=1}^{m_0+1} \sum_{j_1, l_1=1}^{m_1} G_{l_0 l_1}^{(1)} \cdot \left(S^{(1)} \right)_{l_0 l_1}^{j_0 j_1} G_{j_0 j_1}^{(1)} \right).$$

Найдем, при каком значении $\alpha^{(1)}$ полученное значение ошибки сети будет минимальным. Для этого решим уравнение

$$\frac{\partial E_s(t+1)}{\partial \alpha^{(1)}} = -\frac{1}{L^2} \cdot \sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \left(G_{j_0 j_1}^{(1)} \right)^2 +$$

$$+ \alpha^{(1)} \cdot \frac{1}{L^3} \cdot \left(\sum_{j_0, l_0=1}^{m_0+1} \sum_{j_1, l_1=1}^{m_1} G_{l_0 l_1}^{(1)} \cdot \left(S^{(1)} \right)_{l_0 l_1}^{j_0 j_1} G_{j_0 j_1}^{(1)} \right) = 0.$$

Полученное значение $\alpha^{(1)}$ равно

$$\alpha^{(1)} = L\cdot \frac{\sum\limits_{j_0=1}^{m_0+1}\sum\limits_{j_1=1}^{m_1}\left(G^{(1)}_{j_0j_1}\right)^2}{\sum\limits_{j_0,l_0=1}^{m_0+1}\sum\limits_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0l_1}\cdot\left(S^{(1)}\right)^{j_0j_1}_{l_0l_1} G^{(1)}_{j_0j_1}}.$$

Таким образом

$$\alpha^{(1)}_m = \alpha^{(1)}\cdot\frac{1}{L} = \frac{\sum\limits_{j_0=1}^{m_0+1}\sum\limits_{j_1=1}^{m_1}\left(G^{(1)}_{j_0j_1}\right)^2}{\sum\limits_{j_0,l_0=1}^{m_0+1}\sum\limits_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0l_1}\cdot\left(S^{(1)}\right)^{j_0j_1}_{l_0l_1} G^{(1)}_{j_0j_1}}.$$

Теорема доказана.

Теорема 2. При двухпараметрическом обучении двухслойной гетерогенной нейронной сети адаптивный шаг для каждого слоя определяется в соответствии с формулами:

$$\alpha^{(1)}_m = \frac{D_1C - D_2B}{AC - B^2},\quad \alpha^{(2)}_m = \frac{D_2A - D_1B}{AC - B^2}.$$

где

$$A = \sum_{j_0,l_0=1}^{m_0+1}\sum_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0l_1}\cdot\left(S^{(1)}\right)^{j_0j_1}_{l_0l_1} G^{(1)}_{j_0j_1},$$

$$B = \sum_{j_0=1}^{m_0+1}\sum_{j_1=1}^{m_1}\sum_{l_1=1}^{m_1+1}\sum_{l_2=1}^{m_2} G^{(2)}_{l_1l_2}\cdot\left(S^{(1,2)}\right)^{j_0j_1}_{l_1l_2}\cdot G^{(1)}_{j_0j_1},$$

$$C = \sum_{j_1,l_1=1}^{m_1+1}\sum_{j_2,l_2=1}^{m_2} G^{(2)}_{l_1l_2}\cdot\left(S^{(2)}\right)^{j_1j_2}_{l_1l_2} G^{(2)}_{j_1j_2},$$

$$D_1 = \sum_{j_0=1}^{m_0+1}\sum_{j_1=1}^{m_1}\left(G^{(1)}_{j_1j_1}\right)^2,\quad D_2 = \sum_{j_1=1}^{m_1+1}\sum_{j_2=1}^{m_2}\left(G^{(2)}_{j_1j_2}\right)^2,$$

При этом модификация синаптических связей происходит в соответствии с формулами:

$$w^{(2)}_{j_1j_2}(t+1) = w^{(2)}_{j_1j_2}(t) - \alpha^{(2)}_m\cdot G^{(2)}_{j_1j_2},\ j_1 = \overline{1,m_1},\ j_2 = \overline{1,m_2}$$

$$T^{(2)}_{j_2}(t+1) = T^{(2)}_{j_2}(t) - \alpha^{(2)}_m\cdot G^{(2)}_{(m_1+1)j_2},\quad j_2 = \overline{1,m_2}$$

$$w^{(1)}_{j_0j_1}(t+1) = w^{(1)}_{j_0j_1}(t) - \alpha^{(1)}_m\cdot G^{(1)}_{j_0j_1},\ j_0 = \overline{1,m_0},\ j_1 = \overline{1,m_1}$$

$$T^{(1)}_{j_1}(t+1) = T^{(1)}_{j_1}(t) - \alpha^{(1)}_m\cdot G^{(1)}_{(m_0+1)j_1},\quad j_1 = \overline{1,m_1},$$

Доказательство.

При модификации синаптических связей, ошибка сети изменяется по формуле

$$E_s(t+1) = \frac{1}{L}\cdot\sum_{k=1}^{L}E_s^{(k)}(t+1) = \frac{1}{L}\sum_{k=1}^{L}E_s^{(k)}(t) +$$

$$+\frac{1}{L}\cdot\left(\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial w^{(2)}_{j_1j_2}}\right)\cdot\left(w^{(2)}_{j_1j_2}(t+1) - w^{(2)}_{j_1j_2}(t)\right) +\right.$$

$$\left.+\sum_{j_2=1}^{m_2}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial T^{(2)}_{j_2}}\right)\cdot\left(T^{(2)}_{j_2}(t+1) - T^{(2)}_{j_2}(t)\right)\right) +$$

$$+\frac{1}{L}\cdot\left(\sum_{j_0=1}^{m_0}\sum_{j_1=1}^{m_1}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial w^{(1)}_{j_0j_1}}\right)\cdot\left(w^{(1)}_{j_0j_1}(t+1) - w^{(1)}_{j_0j_1}(t)\right) +\right.$$

$$\left.+\sum_{j_1=1}^{m_1}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial T^{(1)}_{j_1}}\right)\cdot\left(T^{(1)}_{j_1}(t+1) - T^{(1)}_{j_1}(t)\right)\right) +$$

$$+\frac{1}{2L}\cdot\sum_{k=1}^{L}\left(\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_1=1}^{m_1}\sum_{l_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial w^{(2)}_{j_1j_2}\partial w^{(2)}_{l_1l_2}}\cdot\left(w^{(2)}_{j_1j_2}(t+1) -\right.\right.$$

$$\left.-w^{(2)}_{j_1j_2}(t)\right)\cdot\left(w^{(2)}_{l_1l_2}(t+1) - w^{(2)}_{l_1l_2}(t)\right)+$$

$$+\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial w^{(2)}_{j_1j_2}\partial T^{(2)}_{l_2}}\cdot\left(w^{(2)}_{j_1j_2}(t+1) - w^{(2)}_{j_1j_2}(t)\right)\times$$

$$\times\left(T^{(2)}_{l_2}(t+1) - T^{(2)}_{l_2}(t)\right)+$$

$$+\sum_{l_1=1}^{m_1}\sum_{l_2=1}^{m_2}\sum_{j_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial T^{(2)}_{j_2}\partial w^{(2)}_{l_1l_2}}\cdot\left(T^{(2)}_{j_2}(t+1) - T^{(2)}_{j_2}(t)\right)\times$$

$$\times\left(w^{(2)}_{l_1l_2}(t+1) - w^{(2)}_{l_1l_2}(t)\right)+$$

$$+\sum_{l_2=1}^{m_2}\sum_{j_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial T^{(2)}_{l_2}\partial T^{(2)}_{j_2}}\cdot\left(T^{(2)}_{l_2}(t+1) - T^{(2)}_{l_2}(t)\right)\times$$

$$\left.\times\left(T^{(2)}_{j_2}(t+1) - T^{(2)}_{j_2}(t)\right)\right)+$$

$$+\frac{1}{2L}\cdot\sum_{k=1}^{L}\left(\sum_{j_0=1}^{m_0}\sum_{j_1=1}^{m_1}\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w^{(1)}_{j_0j_1}\partial w^{(1)}_{l_0l_1}}\cdot\left(w^{(1)}_{j_0j_1}(t+1) -\right.\right.$$

$$\left.-w^{(1)}_{j_0j_1}(t)\right)\cdot\left(w^{(1)}_{l_0l_1}(t+1) - w^{(1)}_{l_0l_1}(t)\right)+$$

$$+\sum_{j_0=1}^{m_0}\sum_{j_1=1}^{m_1}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w^{(1)}_{j_0j_1}\partial T^{(1)}_{l_1}}\cdot\left(w^{(1)}_{j_0j_1}(t+1) - w^{(1)}_{j_0j_1}(t)\right)\times$$

$$\times\left(T^{(1)}_{l_1}(t+1) - T^{(1)}_{l_1}(t)\right)+$$

$$+\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\sum_{j_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial T^{(1)}_{j_1}\partial w^{(1)}_{l_0l_1}}\cdot\left(T^{(1)}_{j_1}(t+1) - T^{(1)}_{j_1}(t)\right)\times$$

$$\times\left(w^{(1)}_{l_0l_1}(t+1) - w^{(1)}_{l_0l_1}(t)\right)+$$

$$+\sum_{l_1=1}^{m_1}\sum_{j_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial T^{(1)}_{l_1}\partial T^{(1)}_{j_1}}\cdot\left(T^{(1)}_{l_1}(t+1) - T^{(1)}_{l_1}(t)\right)\times$$

$$\left.\times\left(T^{(1)}_{j_1}(t+1) - T^{(1)}_{j_1}(t)\right)\right)+$$

$$+2\cdot\frac{1}{2L}\cdot\sum_{k=1}^{L}\left(\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w^{(2)}_{j_1j_2}\partial w^{(1)}_{l_0l_1}}\cdot\left(w^{(2)}_{j_1j_2}(t+1) -\right.\right.$$

$$\left.-w^{(2)}_{j_1j_2}(t)\right)\cdot\left(w^{(1)}_{l_0l_1}(t+1) - w^{(1)}_{l_0l_1}(t)\right)+$$

$$+\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w^{(2)}_{j_1j_2}\partial T^{(1)}_{l_1}}\cdot\left(w^{(2)}_{j_1j_2}(t+1) - w^{(2)}_{j_1j_2}(t)\right)\times$$

$$\times\left(T^{(1)}_{l_1}(t+1) - T^{(1)}_{l_1}(t)\right)+$$

$$+\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\sum_{j_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial T^{(2)}_{j_2}\partial w^{(1)}_{l_0l_1}}\cdot\left(T^{(2)}_{j_2}(t+1) - T^{(2)}_{j_2}(t)\right)\times$$

$$\times\left(w^{(1)}_{l_0l_1}(t+1) - w^{(1)}_{l_0l_1}(t)\right)+$$

$$+\sum_{l_2=1}^{m_2}\sum_{j_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial T^{(2)}_{l_2}\partial T^{(1)}_{j_1}}\cdot\left(T^{(2)}_{l_2}(t+1) - T^{(2)}_{l_2}(t)\right)\times$$

$$\left.\times\left(T^{(1)}_{j_1}(t+1) - T^{(1)}_{j_1}(t)\right)\right) =$$

$$= E_s(t) - \alpha^{(2)} \cdot \frac{1}{L^2} \cdot \sum_{j_1=1}^{m_1+1} \sum_{j_2=1}^{m_2} \left(G^{(2)}_{j_1 j_2} \right)^2 - \alpha^{(1)} \cdot \frac{1}{L^2} \cdot \sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \left(G^{(1)}_{j_1 j_1} \right)^2 +$$

$$+ \left(\alpha^{(2)} \right)^2 \cdot \frac{1}{2L^3} \cdot \left(\sum_{j_1,l_1=1}^{m_1+1} \sum_{j_2,l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(2)} \right)^{j_1 j_2}_{l_1 l_2} \cdot G^{(2)}_{j_1 j_2} \right) +$$

$$+ \left(\alpha^{(1)} \right)^2 \cdot \frac{1}{2L^3} \cdot \left(\sum_{j_0,l_0=1}^{m_0+1} \sum_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0 l_1} \cdot \left(S^{(1)} \right)^{j_0 j_1}_{l_0 l_1} \cdot G^{(1)}_{j_0 j_1} \right) +$$

$$+ \alpha^{(1)} \cdot \alpha^{(2)} \cdot \frac{1}{L^3} \cdot \left(\sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \sum_{l_1=1}^{m_1+1} \sum_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right)$$

Найдем минимум функции ошибки, зависящей от двух переменных $\alpha^{(1)}$ и $\alpha^{(2)}$. Для этого найдем стационарную точку, приравняв частные производные к нулю. Данная точка и будет являться точкой минимума.

$$\frac{\partial E_s(t+1)}{\partial \alpha^{(1)}} = -\frac{1}{L^2} \cdot \sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \left(G^{(1)}_{j_1 j_1} \right)^2 +$$

$$+ \alpha^{(1)} \cdot \frac{1}{L^3} \cdot \left(\sum_{j_0,l_0=1}^{m_0+1} \sum_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0 l_1} \cdot \left(S^{(1)} \right)^{j_0 j_1}_{l_0 l_1} \cdot G^{(1)}_{j_0 j_1} \right) +$$

$$+ \alpha^{(2)} \cdot \frac{1}{L^3} \cdot \left(\sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \sum_{l_1=1}^{m_1+1} \sum_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right),$$

$$\frac{\partial E_s(t+1)}{\partial \alpha^{(2)}} = -\frac{1}{L^2} \cdot \sum_{j_1=1}^{m_1+1} \sum_{j_2=1}^{m_2} \left(G^{(2)}_{j_1 j_2} \right)^2 +$$

$$+ \alpha^{(2)} \cdot \frac{1}{L^3} \cdot \left(\sum_{j_1,l_1=1}^{m_1+1} \sum_{j_2,l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(2)} \right)^{j_1 j_2}_{l_1 l_2} \cdot G^{(2)}_{j_1 j_2} \right) +$$

$$+ \alpha^{(1)} \cdot \frac{1}{L^3} \cdot \left(\sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \sum_{l_1=1}^{m_1+1} \sum_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right).$$

То есть необходимо решить систему уравнений:

$$\alpha^{(1)} \cdot \left(\sum_{j_0,l_0=1}^{m_0+1} \sum_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0 l_1} \cdot \left(S^{(1)} \right)^{j_0 j_1}_{l_0 l_1} \cdot G^{(1)}_{j_0 j_1} \right) +$$

$$+ \alpha^{(2)} \cdot \left(\sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \sum_{l_1=1}^{m_1+1} \sum_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) = L \cdot \sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \left(G^{(1)}_{j_1 j_1} \right)^2$$

$$\alpha^{(1)} \cdot \left(\sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \sum_{l_1=1}^{m_1+1} \sum_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) +$$

$$+ \alpha^{(2)} \cdot \left(\sum_{j_1,l_1=1}^{m_1+1} \sum_{j_2,l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(2)} \right)^{j_1 j_2}_{l_1 l_2} \cdot G^{(2)}_{j_1 j_2} \right) = L \cdot \sum_{j_1=1}^{m_1+1} \sum_{j_2=1}^{m_2} \left(G^{(2)}_{j_1 j_2} \right)^2$$

Тогда

$$\alpha^{(1)}_m = \alpha^{(1)} \cdot \frac{1}{L} = \frac{\begin{vmatrix} D_1 & B \\ D_2 & C \end{vmatrix}}{\begin{vmatrix} A & B \\ B & C \end{vmatrix}} = \frac{D_1 C - D_2 B}{AC - B^2},$$

$$\alpha^{(2)}_m = \alpha^{(2)} \cdot \frac{1}{L} = \frac{\begin{vmatrix} A & D_1 \\ B & D_2 \end{vmatrix}}{\begin{vmatrix} A & B \\ B & C \end{vmatrix}} = \frac{D_2 A - D_1 B}{AC - B^2}.$$

Теорема доказана.

Теорема 3. При обучении обобщенным методом наискорейшего спуска двухслойной гетерогенной нейронной сети

В соответствии с методом Крамера получаем решение:

$$\alpha^{(1)} = \frac{\begin{vmatrix} L \cdot \sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \left(G^{(1)}_{j_1 j_1} \right)^2 & \left(\sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \sum\limits_{l_1=1}^{m_1+1} \sum\limits_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) \\ L \cdot \sum\limits_{j_1=1}^{m_1+1} \sum\limits_{j_2=1}^{m_2} \left(G^{(2)}_{j_1 j_2} \right)^2 & \left(\sum\limits_{j_1,l_1=1}^{m_1+1} \sum\limits_{j_2,l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(2)} \right)^{j_1 j_2}_{l_1 l_2} \cdot G^{(2)}_{j_1 j_2} \right) \end{vmatrix}}{\begin{vmatrix} \left(\sum\limits_{j_0,l_0=1}^{m_0+1} \sum\limits_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0 l_1} \cdot \left(S^{(1)} \right)^{j_0 j_1}_{l_0 l_1} \cdot G^{(1)}_{j_0 j_1} \right) & \left(\sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \sum\limits_{l_1=1}^{m_1+1} \sum\limits_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) \\ \left(\sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \sum\limits_{l_1=1}^{m_1+1} \sum\limits_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) & \left(\sum\limits_{j_1,l_1=1}^{m_1+1} \sum\limits_{j_2,l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(2)} \right)^{j_1 j_2}_{l_1 l_2} \cdot G^{(2)}_{j_1 j_2} \right) \end{vmatrix}}$$

$$\alpha^{(2)} = \frac{\begin{vmatrix} \left(\sum\limits_{j_0,l_0=1}^{m_0+1} \sum\limits_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0 l_1} \cdot \left(S^{(1)} \right)^{j_0 j_1}_{l_0 l_1} \cdot G^{(1)}_{j_0 j_1} \right) & L \cdot \sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \left(G^{(1)}_{j_1 j_1} \right)^2 \\ \left(\sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \sum\limits_{l_1=1}^{m_1+1} \sum\limits_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) & L \cdot \sum\limits_{j_1=1}^{m_1+1} \sum\limits_{j_2=1}^{m_2} \left(G^{(2)}_{j_1 j_2} \right)^2 \end{vmatrix}}{\begin{vmatrix} \left(\sum\limits_{j_0,l_0=1}^{m_0+1} \sum\limits_{j_1,l_1=1}^{m_1} G^{(1)}_{l_0 l_1} \cdot \left(S^{(1)} \right)^{j_0 j_1}_{l_0 l_1} \cdot G^{(1)}_{j_0 j_1} \right) & \left(\sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \sum\limits_{l_1=1}^{m_1+1} \sum\limits_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) \\ \left(\sum\limits_{j_0=1}^{m_0+1} \sum\limits_{j_1=1}^{m_1} \sum\limits_{l_1=1}^{m_1+1} \sum\limits_{l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(1,2)} \right)^{j_0 j_1}_{l_1 l_2} \cdot G^{(1)}_{j_0 j_1} \right) & \left(\sum\limits_{j_1,l_1=1}^{m_1+1} \sum\limits_{j_2,l_2=1}^{m_2} G^{(2)}_{l_1 l_2} \cdot \left(S^{(2)} \right)^{j_1 j_2}_{l_1 l_2} \cdot G^{(2)}_{j_1 j_2} \right) \end{vmatrix}}$$

адаптивный шаг для каждого слоя определяется в соответствии с формулами

$$\alpha_m = \frac{D_1 + D_2}{A + 2B + C}\,.$$

где A, B, C, D_1 и D_2 определены также, как в теореме 2.

При этом модификация синаптических связей происходит в соответствии с формулами

$$w_{j_1 j_2}^{(2)}(t+1) = w_{j_1 j_2}^{(2)}(t) - \alpha_m \cdot G_{j_1 j_2}^{(2)},\ j_1 = \overline{1, m_1},\ j_2 = \overline{1, m_2}$$

$$T_{j_2}^{(2)}(t+1) = T_{j_2}^{(2)}(t) - \alpha_m \cdot G_{(m_1+1) j_2}^{(2)}, \quad j_2 = \overline{1, m_2}$$

$$w_{j_0 j_1}^{(1)}(t+1) = w_{j_0 j_1}^{(1)}(t) - \alpha_m \cdot G_{j_0 j_1}^{(1)},\ j_0 = \overline{1, m_0},\ j_1 = \overline{1, m_1}$$

$$T_{j_1}^{(1)}(t+1) = T_{j_1}^{(1)}(t) - \alpha_m \cdot G_{(m_0+1) j_1}^{(1)}, \quad j_1 = \overline{1, m_1}\ ,$$

Доказательство.

При модификации синаптических связей, ошибка сети изменяется по формуле

$$E_s(t+1) = \frac{1}{L}\cdot\sum_{k=1}^{L} E_s^{(k)}(t+1) = \frac{1}{L}\sum_{k=1}^{L} E_s^{(k)}(t) +$$

$$+\frac{1}{L}\cdot\left(\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}}\right)\cdot\left(w_{j_1 j_2}^{(2)}(t+1) - w_{j_1 j_2}^{(2)}(t)\right) +\right.$$

$$\left.+\sum_{j_2=1}^{m_2}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial T_{j_2}^{(2)}}\right)\cdot\left(T_{j_2}^{(2)}(t+1) - T_{j_2}^{(2)}(t)\right)\right) +$$

$$+\frac{1}{L}\cdot\left(\sum_{j_0=1}^{m_0}\sum_{j_1=1}^{m_1}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial w_{j_0 j_1}^{(1)}}\right)\cdot\left(w_{j_0 j_1}^{(1)}(t+1) - w_{j_0 j_1}^{(1)}(t)\right) +\right.$$

$$\left.\sum_{j_1=1}^{m_1}\left(\sum_{k=1}^{L}\frac{\partial E_s^{(k)}}{\partial T_{j_1}^{(1)}}\right)\cdot\left(T_{j_1}^{(1)}(t+1) - T_{j_1}^{(1)}(t)\right) +\right) +$$

$$+\frac{1}{2L}\cdot\sum_{k=1}^{L}\left(\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_1=1}^{m_1}\sum_{l_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}\partial w_{l_1 l_2}^{(2)}}\cdot\left(w_{j_1 j_2}^{(2)}(t+1) -\right.\right.$$

$$\left.-w_{j_1 j_2}^{(2)}(t)\right)\cdot\left(w_{l_1 l_2}^{(2)}(t+1) - w_{l_1 l_2}^{(2)}(t)\right) +$$

$$+\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}\partial T_{l_2}^{(2)}}\cdot\left(w_{j_1 j_2}^{(2)}(t+1) - w_{j_1 j_2}^{(2)}(t)\right)\times$$

$$\times\left(T_{l_2}^{(2)}(t+1) - T_{l_2}^{(2)}(t)\right) +$$

$$+\sum_{l_1=1}^{m_1}\sum_{l_2=1}^{m_2}\sum_{j_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial T_{j_2}^{(2)}\partial w_{l_1 l_2}^{(2)}}\cdot\left(T_{j_2}^{(2)}(t+1) - T_{j_2}^{(2)}(t)\right)\times$$

$$\times\left(w_{l_1 l_2}^{(2)}(t+1) - w_{l_1 l_2}^{(2)}(t)\right) +$$

$$+\sum_{l_2=1}^{m_2}\sum_{j_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial T_{l_2}^{(2)}\partial T_{j_2}^{(2)}}\cdot\left(T_{l_2}^{(2)}(t+1) - T_{l_2}^{(2)}(t)\right)\times$$

$$\left.\times\left(T_{j_2}^{(2)}(t+1) - T_{j_2}^{(2)}(t)\right)\right) +$$

$$+\frac{1}{2L}\cdot\sum_{k=1}^{L}\left(\sum_{j_0=1}^{m_0}\sum_{j_1=1}^{m_1}\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w_{j_0 j_1}^{(1)}\partial w_{l_0 l_1}^{(1)}}\cdot\left(w_{j_0 j_1}^{(1)}(t+1) -\right.\right.$$

$$\left.-w_{j_0 j_1}^{(1)}(t)\right)\cdot\left(w_{l_0 l_1}^{(1)}(t+1) - w_{l_0 l_1}^{(1)}(t)\right) +$$

$$+\sum_{j_0=1}^{m_0}\sum_{j_1=1}^{m_1}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w_{j_0 j_1}^{(1)}\partial T_{l_1}^{(1)}}\cdot\left(w_{j_0 j_1}^{(1)}(t+1) - w_{j_0 j_1}^{(1)}(t)\right)\times$$

$$\times\left(T_{l_1}^{(1)}(t+1) - T_{l_1}^{(1)}(t)\right) +$$

$$+\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\sum_{j_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial T_{j_1}^{(1)}\partial w_{l_0 l_1}^{(1)}}\cdot\left(T_{j_1}^{(1)}(t+1) - T_{j_1}^{(1)}(t)\right)\times$$

$$\times\left(w_{l_0 l_1}^{(1)}(t+1) - w_{l_0 l_1}^{(1)}(t)\right) +$$

$$+\sum_{l_1=1}^{m_1}\sum_{j_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial T_{l_1}^{(1)}\partial T_{j_1}^{(1)}}\cdot\left(T_{l_1}^{(1)}(t+1) - T_{l_1}^{(1)}(t)\right)\times$$

$$\left.\times\left(T_{j_1}^{(1)}(t+1) - T_{j_1}^{(1)}(t)\right)\right) +$$

$$+2\cdot\frac{1}{2L}\cdot\sum_{k=1}^{L}\left(\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}\partial w_{l_0 l_1}^{(1)}}\cdot\left(w_{j_1 j_2}^{(2)}(t+1) -\right.\right.$$

$$\left.-w_{j_1 j_2}^{(2)}(t)\right)\cdot\left(w_{l_0 l_1}^{(1)}(t+1) - w_{l_0 l_1}^{(1)}(t)\right) +$$

$$+\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2}\sum_{l_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}\partial T_{l_1}^{(1)}}\cdot\left(w_{j_1 j_2}^{(2)}(t+1) - w_{j_1 j_2}^{(2)}(t)\right)\times$$

$$\times\left(T_{l_1}^{(1)}(t+1) - T_{l_1}^{(1)}(t)\right) +$$

$$+\sum_{l_0=1}^{m_0}\sum_{l_1=1}^{m_1}\sum_{j_2=1}^{m_2}\frac{\partial^2 E_s^{(k)}}{\partial T_{j_2}^{(2)}\partial w_{l_0 l_1}^{(1)}}\cdot\left(T_{j_2}^{(2)}(t+1) - T_{j_2}^{(2)}(t)\right)\times$$

$$\times\left(w_{l_0 l_1}^{(1)}(t+1) - w_{l_0 l_1}^{(1)}(t)\right) +$$

$$+\sum_{l_2=1}^{m_2}\sum_{j_1=1}^{m_1}\frac{\partial^2 E_s^{(k)}}{\partial T_{l_2}^{(2)}\partial T_{j_1}^{(1)}}\cdot\left(T_{l_2}^{(2)}(t+1) - T_{l_2}^{(2)}(t)\right)\times$$

$$\left.\times\left(T_{j_1}^{(1)}(t+1) - T_{j_1}^{(1)}(t)\right)\right) =$$

$$= E_s(t) - \alpha\cdot\frac{1}{L^2}\cdot\sum_{j_1=1}^{m_1+1}\sum_{j_2=1}^{m_2}\left(G_{j_1 j_2}^{(2)}\right)^2 - \alpha\cdot\frac{1}{L^2}\cdot\sum_{j_0=1}^{m_0+1}\sum_{j_1=1}^{m_1}\left(G_{j_1 j_1}^{(1)}\right)^2 +$$

$$+\alpha^2\cdot\frac{1}{2L^3}\cdot\left(\sum_{j_1,l_1=1}^{m_1+1}\sum_{j_2,l_2=1}^{m_2} G_{l_1 l_2}^{(2)}\cdot\left(S^{(2)}\right)_{l_1 l_2}^{j_1 j_2}\cdot G_{j_1 j_2}^{(2)}\right) +$$

$$+\alpha^2\cdot\frac{1}{2L^3}\cdot\left(\sum_{j_0,l_0=1}^{m_0+1}\sum_{j_1,l_1=1}^{m_1} G_{l_0 l_1}^{(1)}\cdot\left(S^{(1)}\right)_{l_0 l_1}^{j_0 j_1}\cdot G_{j_0 j_1}^{(1)}\right) +$$

$$+\alpha^2\cdot\frac{1}{L^3}\cdot\left(\sum_{j_0=1}^{m_0+1}\sum_{j_1=1}^{m_1}\sum_{l_1=1}^{m_1+1}\sum_{l_2=1}^{m_2} G_{l_1 l_2}^{(2)}\cdot\left(S^{(1,2)}\right)_{l_1 l_2}^{j_0 j_1}\cdot G_{j_0 j_1}^{(1)}\right).$$

Найдем значение параметра обучения α, при котором функция ошибки обучения достигает наименьшего значения. Для этого необходимо решить уравнение

$$\frac{\partial E_s(t+1)}{\partial\alpha} = -\frac{1}{L^2}\cdot\sum_{j_1=1}^{m_1+1}\sum_{j_2=1}^{m_2}\left(G_{j_1 j_2}^{(2)}\right)^2 - \frac{1}{L^2}\cdot\sum_{j_0=1}^{m_0+1}\sum_{j_1=1}^{m_1}\left(G_{j_1 j_1}^{(1)}\right)^2 +$$

$$+\alpha\cdot\frac{1}{L^3}\cdot\left(\sum_{j_1,l_1=1}^{m_1+1}\sum_{j_2,l_2=1}^{m_2} G_{l_1 l_2}^{(2)}\cdot\left(S^{(2)}\right)_{l_1 l_2}^{j_1 j_2}\cdot G_{j_1 j_2}^{(2)}\right) +$$

$$+\alpha\cdot\frac{1}{L^3}\cdot\left(\sum_{j_0,l_0=1}^{m_0+1}\sum_{j_1,l_1=1}^{m_1} G_{l_0 l_1}^{(1)}\cdot\left(S^{(1)}\right)_{l_0 l_1}^{j_0 j_1}\cdot G_{j_0 j_1}^{(1)}\right) +$$

$$+2\boldsymbol{\alpha}\cdot\frac{1}{\boldsymbol{L}^3}\cdot\left(\sum_{j_0=1}^{m_0+1}\sum_{j_1=1}^{m_1}\sum_{l_1=1}^{m_1+1}\sum_{l_2=1}^{m_2}\boldsymbol{G}^{(2)}_{l_1l_2}\cdot\left(\boldsymbol{S}^{(1,2)}\right)^{j_0j_1}_{l_1l_2}\cdot\boldsymbol{G}^{(1)}_{j_0j_1}\right)=0\,.$$

Тогда шаг обучения $\boldsymbol{\alpha}$ определяется в соответствии с выражением:

$$\boldsymbol{\alpha}=\boldsymbol{L}\cdot\frac{\boldsymbol{D}_1+\boldsymbol{D}_2}{\boldsymbol{A}+2\boldsymbol{B}+\boldsymbol{C}}\,,$$

А значит

$$\boldsymbol{\alpha}_m=\boldsymbol{\alpha}\cdot\frac{1}{\boldsymbol{L}}=\frac{\boldsymbol{D}_1+\boldsymbol{D}_2}{\boldsymbol{A}+2\boldsymbol{B}+\boldsymbol{C}}$$

Теорема доказана.

4. ЗАКЛЮЧЕНИЕ

В данной статье предложена методика обучения двухслойных гетерогенных сетей прямого распространения на основе метода наискорейшего спуска: послойное обучение, двухпараметрическое обучение и обобщенный метод наискорейшего спуска. Для каждого из этих подходов приведены и доказаны соотношения для изменения синаптических связей и формулы для вычисления адаптивных шагов. Эксперименты, проведенные на тестовых данных, показывают более быструю сходимость метода двухпараметрического обучения, хотя каждый из выше предложенных методов может дать лучший результат в сравнении с двумя другими, при определенных тестовых данных.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Danuta Rutkowska, Maciej Piliński, Leszek Rutkowski. Sieci neuronowe, algorytmy geneticzne I systemy rozmyte. – Warszawa, Łódź: Wydawnictwo Naukowe PWN, 1999. – 412 s.
2. Hornik K., Stinchcombe M. and White H. Multilayer feedforward networks are universal approximators // Neural Networks, 1989, vol.2 – PP. 359-366.
3. Гладкий И.И., Головко В.А., Махнист Л.П. Обучение нейронных сетей с использованием метода наискорейшего спуска // Вестник Брестского государственного технического университета. – Брест: БГТУ, 2001. № 5: Физика, математика, химия. – С. 56-61.
4. Маньяков Н.В., Махнист Л.П. Матричная алгоритмизация обучения многослойных нейронных сетей с использованием градиентных методов // Вестник Брестского государственного технического университета. – Брест: БГТУ, 2002. - №5 (17): Физика, математика, химия. – С. 60-64.
5. V. Golovko, N. Maniakov, L. Makhnist. Multilayer Neural Networks Training Methodic // Proceedings of the Workshop Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS’2003). – Lviv, Ukraine, September 8-10, 2003. – PP. 185-190.

УДК 681.3

Головко В.А., Дунец А.П., Кирись А.Н., Селезнев П.В.

РАСПОЗНАВАНИЕ РЕГИСТРАЦИОННЫХ НОМЕРОВ АВТОМОБИЛЕЙ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

Введение

Автоматическое распознавание регистрационных номеров автомобилей в реальном режиме времени является актуальной задачей, решение которой может использоваться во многих практических приложениях. Это, например, контроль автомобилей на автострадах или при парковке на автостоянках. При этом предполагается, что имеется видеокамера, которая отображает изображение регистрационного номера автомобиля. Изображение является зашумленным. Задачей системы распознавания является корректная идентификация номера автомобиля в реальном масштабе времени и представление его в текстовом формате.

В настоящее время имеется много научных публикаций и коммерческого программного обеспечения, предназначенного для распознавания регистрационных номеров автомобилей [1,2,3]. Основная проблема здесь заключается в повышении качества распознавания.

В данной статье описываются и анализируются два нейросетевых метода распознавания автомобильных номеров. Первый метод базируется на преобразовании Фурье исходного изображения номера автомобиля и применении многослойного персептрона для распознавания. Второй подход основывается на применении в качестве распознавателя самоорганизующейся сети Кохонена. Результаты экспериментов обсуждаются.

1. Структура системы распознавания

Рассмотрим общую структуру системы распознавания регистрационных номеров автомобилей. Она состоит из модуля сегментации, модуля обработки изображений и модуля нейронный сетей (рис.1).

Модуль сегментации осуществляет разделение исходного изображения номера автомобиля на отдельные участки, каждый из которых соответствует определённому знаку (букве или цифре) регистрационного номера (рис.2). Для этого можно использовать градиентный метод, описанный в [4].

Модуль обработки производит преобразование изображения номера автомобиля в форму удобную для подачи на модуль нейронных сетей. В данной работе рассмотрено два варианта реализации данного модуля.

Рис.1. Структурная схема системы распознавания.

Дунец Андрей Петрович, *ст. преподаватель каф. интеллектуальных информационных технологий Брестского государственного технического университета.*
Кирись А.Н, *ст. преподаватель каф. ИИТ Брестского государственного технического университета.*
Селезнев Петр Владимирович, *ст. преподаватель каф. ИИТ Брестского государственного технического университета.*
Беларусь, БГТУ, 224017, г. Брест, ул. Московская, 267.