УДК: 550.837

ПРИМЕНЕНИЕ НЕЙРОСЕТЕЙ СВЕРТОЧНОГО ТИПА В ОБРАТНЫХ ЗАДАЧАХ ГЕОЭЛЕКТРИКИ

© 2025 г. М. И. Шимелевич[1, *], Е. А. Родионов[1, **], И. Е. Оборнев[1, 2, *], Е. А. Оборнев[1, ****]**

[1]*Российский государственный геологоразведочный университет им. Серго Орджоникидзе, г. Москва, Россия*
[2]*Научно-исследовательский институт ядерной физики им. Д.В. Скобельцына МГУ им. М.В. Ломоносова, г. Москва, Россия*
**E-mail*: *shimelevich-m@yandex.ru*
*** E-mail*: *evgeny_980@list.ru*
**** E-mail*: *o_ivano@mail.ru*
***** E-mail*: *obornevea@mail.ru*

Поступила в редакцию 25.11.2023 г.
После доработки 24.06.2024 г.
Принята к публикации 25.06.24 г.

В настоящее время нейронные сети успешно применяются для решения обратных и других задач геофизики. Цель данной работы, которая является продолжением серии работ авторов, состоит в повышении эффективности НС метода решения нелинейных обратных 3D-задач геоэлектрики на основе конструирования авторской нейронной сети сверточного типа. Сеть включает ряд дополнительных специальных преобразований (сжатие данных, подавление влияния неизвестной фоновой среды и др.), предшествующих обучению классической MLP-нейросети и адаптированных к решаемой обратной задаче. Это позволяет формализовано, исключая человеческий фактор, решать обратные задачи геоэлектрики большой размерности без задания первого приближения на основе данных, измеренных в областях, размеры которых превышают размеры области обучения сети. Скорость инверсии составляет первые десятки секунд и практически не зависит от физической размерности (2D или 3D) данных. Найденное с помощью обученной нейросети решение обратной задачи, при необходимости, может уточняться методом случайного поиска. Приводятся численные результаты решения 3D-задач геоэлектрики на модельных и полевых данных, подтверждающие заявленные параметры разработки.

Ключевые слова: МТЗ, 3D-обратная задача, геоэлектрика, нейронные сети, сверточные сети, НС аппроксиматор, уточнение нейросетевого решения.

DOI: 10.31857/S0002333725010061, **EDN:** ACMGSS

1. ВВЕДЕНИЕ

В 90-х годах прошлого века были получены первые результаты применения нейросетевого (НС) метода *решения нелинейных* (*коэффициентных*) *обратных задач* геоэлектрики [Poulton et al., 1989; Raiche, 1991; Poulton et al., 1992; Hidalgo et al., 1994; Спичак, Попова, 1998; Shimelevich, Obornev, 1998]. В этих работах использовались стандартные алгоритмы обучения “с учителем” простейших полносвязных трехслойных сетей (типа персептрон) на обучающем множестве опорных решений прямых (а значит, и обратных) задач, которое строится заранее с помощью прямого оператора задачи.

В работе [Шимелевич, Оборнев, 1999] было показано, что математический смысл НС метода, применяемого в указанных работах, заключается в аппроксимации обратного оператора условно-корректной обратной задачи с помощью нейросетевой аппроксимационно-интерполяционной конструкции, которая проводится наилучшим образом через точки множества опорных решений. В задачах теории аппроксимации функций такую конструкцию для краткости называют НС аппроксиматором

[Ротштейн, Митюшкин, 1998]. Алгоритмическая идея решения обратной задачи на основе заранее построенного множества опорных решений, сформулированная в общем виде в работе [Тихонов и др., 1983], издавна использовалась геофизиками в палеточном методе, обучающее множество опорных решений при этом именовалось набором палеток. Сопоставление измеренных данных с данными, имеющимися в банке решений, интерполяция и нахождение интерполяционного решения обратной задачи в палеточном методе производилось “вручную”, а в НС методе — в автоматизированном режиме с помощью обученного НС аппроксиматора.

В геофизической практике широко применяются различные методы решения целого ряда задач обработки и интерпретации измеряемых данных с использованием нейросетей под общим названием “НС методы”: построение различных зависимостей, восполнение пропущенных данных [Цымблер, Юртин, 2023], задачи инверсии в условиях недостаточности данных [Спичак, 2019], классификация данных [Шимелевич, Оборнев, 1999; Оборнев и др., 2007; Спичак и др., 2007; Спичак, 2019], электромагнитная томография [Корженевский, 2001; Спичак и др., 2007; 2019; и др.]. Эти методы обычно позиционируются как стохастические, а обучающее множество формируется на основе экспериментальных, литературных, архивных и других типов данных [Спичак, 2019]. НС метод решения обратных задач в контексте работ, упомянутых в начале введения, занимает особое место среди НС методов интерпретации геофизических данных, так как предполагает знание оператора прямой задачи для расчета множества опорных решений и построение приближенного обратного оператора решаемой задачи на основе этого множества. Для того чтобы обозначить специфику НС метода при решении обратных задач и подчеркнуть его аппроксимационную (детерминистскую) сущность, авторами статьи был предложен термин *аппроксимационный нейросетевой* (АНС) метод [Шимелевич, Оборнев, 2009; 2012; Шимелевич и др., 2017].

В работе [Шимелевич и др., 2018] представлена базовая версия алгоритма регуляризованного АНС метода. В последующих работах авторов усовершенствовались отдельные блоки данной версии. В работе [Шимелевич и др., 2022] был представлен авторский алгоритм уточнения НС решения методом случайного поиска с предварительной кластеризацией методом *k*-средних и применением интерполяционных процедур методом IDW-обратных взвешенных расстояний, а также разработана концепция применения глубоких сетей сверточного типа при обучении НС аппроксиматоров. В данной статье представлен модернизированный блок обучения НС аппроксиматора, основу которого составляет авторская версия нейросети сверточного типа. Предлагаемая сеть позволила уменьшить ошибку обучения НС аппроксиматоров по сравнению с известными предыдущими версиями и в целом повысить разрешающую способность АНС метода применительно к решению обратной задачи геоэлектрики.

Приводятся численные примеры решения обратной 3D-задачи геоэлектрики с использованием модельных и полевых наблюдений участка Северо-Пясинской площади.

2. АППРОКСИМАЦИОННЫЙ НС МЕТОД

2.1. Аппроксимационный подход. Понятие аппроксиматора

Обратная коэффициентная условно-корректная задача геоэлектрики рассматривается в классах сред с кусочно-постоянным распределением удельного сопротивления среды, задаваемого на сетке параметризации θ_N размерности N, покрывающей область исследований Ω. После перехода к конечномерной аппроксимации эта задача может быть сведена к решению нелинейного операторного уравнения 1-го рода вида [Тихонов и др., 1983; Дмитриев, 2012; Шимелевич и др., 2017]

$$A_N \boldsymbol{s} = \boldsymbol{f}, \quad \boldsymbol{s} \in S_N \subset R^N, \quad \boldsymbol{f} \in R^M, M \geq N, \tag{1}$$

где $\boldsymbol{s} = (s^1, \ldots, s^N)$ — вектор искомых параметров среды, параметрами задачи являются значения логарифма удельного сопротивления $\lg\rho^n$ в ячейках сетки θ_N:

$$s^n = \lg\rho^n, \; n = 1, \ldots, N;$$

$\boldsymbol{f} = (f^1, \ldots, f^M)$ — вектор входных данных, определяемых на верхней границе Ω_1 области Ω; A_N — непрерывный оператор численного решения прямой задачи; S_N — множество априорных ограничений задачи, определяемое диапазоном D_s возможных изменений логарифма удельного сопротивления реальных сред:

$$S_N: \; [s_{\min} \leq s^n \leq s_{\min} + D_s], \quad n = 1, \ldots, N. \tag{2}$$

Каким бы численным методом не решалось уравнение (1), его приближенное решение $\tilde{s} = (\tilde{s}^1, \ldots, \tilde{s}^N)$ является некоторой векторной функцией (задаваемой алгоритмом решения обратной задачи) $\bar{\Psi} = (\psi^1,...,\psi^N)$ от M переменных — компонент вектора входных данных $\boldsymbol{f} = (f^1,...,f^M)$:

$$\boldsymbol{s} = \bar{\Psi}(\boldsymbol{f}).$$

Аппроксимационный подход к решению уравнения (1) заключается в построении некоторой непрерывной (по $\hat{\boldsymbol{a}}$ и $\boldsymbol{f}$) векторной функции $\Psi(\hat{\boldsymbol{a}},\boldsymbol{f}) = (\psi^1,...,\psi^N)$, задаваемой в аналитическом виде и зависящей от свободных коэффициентов, составляющих некоторую матрицу $\hat{\boldsymbol{a}}$. В качестве функции $\Psi(\hat{\boldsymbol{a}}, \boldsymbol{f})$ может быть задан, например, некоторый полином. Функцию $\Psi(\hat{\boldsymbol{a}},\boldsymbol{f})$ называют аппроксиматором инверсии (или просто аппроксиматором) для уравнения (1). Для определения коэффициентов $\hat{\boldsymbol{a}}$ заранее с помощью прямого оператора A_N рассчитывается банк опорных решений $\mathbb{Q}_{bs}$:

$$\mathbb{Q}_{bs} = \left\{\boldsymbol{s}_i, \boldsymbol{f}_i : \boldsymbol{s}_i \in S_N,\ \boldsymbol{f}_i = A_N \boldsymbol{s}_i,\ i = 1,...,Q_{bs}\right\}.$$

Численное нахождение коэффициентов матрицы $\hat{\boldsymbol{a}}$ (обучение аппроксиматора) сводится к решению оптимизационной задачи вида

$$\sum_{i=1}^{Q_{bs}} \left\|\Psi(\hat{\boldsymbol{a}},\ \boldsymbol{f}_i) - \boldsymbol{s}_i\right\|^2 \longrightarrow \min_{\hat{\boldsymbol{a}}}. \qquad (3)$$

Обученный аппроксиматор является приближенным обратным оператором для задачи (1)–(2). На основе независимой “экзаменационной” выборки известных решений, не использовавшихся при обучении, вычисляется ошибка обучения аппроксиматора, связанная с заменой точного обратного оператора аппроксиматором (собственная ошибка аппроксиматора), которая в общем случае отлична от нуля даже при точных входных данных.

Основное преимущество аппроксимационного подхода состоит в том, что приближенное решение уравнения (1) с помощью обученного аппроксиматора может быть легко и быстро вычислено (как значение функции, определяющей аппроксиматор, заданной в аналитическом виде) для любого вектора входных данных $\boldsymbol{f} \in R^M$ без задания первого приближения

$$\tilde{s} = \Psi(\boldsymbol{f}). \qquad (4)$$

Устойчивость получаемых решений по отношению к малым изменениям входных данных, обеспечивается непрерывностью аппроксиматора по $\boldsymbol{f}$.

2.2. Аппроксимационный нейросетевой (АНС) метод

В АНС методе координатные функции аппроксиматора задаются в виде НС полиномов (нейронных сетей). Простейшей нейронной сетью является трехслойный персептрон с одним скрытым слоем, который описывается следующей формулой

$$\varphi(\boldsymbol{V},\boldsymbol{W},\boldsymbol{x}) = \sum_{l=1}^{L} v_l \lambda\left(\sum_{m=1}^{M} w_{l,m} x_m\right),$$

где λ — функция активации — заданная непрерывная нелинейная монотонно возрастающая ограниченная функция, например, $\lambda(x) = 1/(1+\exp(-x))$ (сигмоид). Вектор $\boldsymbol{V} = (v_1,...,v_L)$ и матрица $\boldsymbol{W}$ определяют матрицу $\hat{\boldsymbol{a}}$ нейросетевого аппроксиматора $\Psi(\hat{\boldsymbol{a}}, \boldsymbol{f})$, а параметр L (количество нейронов на скрытом слое нейронной сети) определяет его сложность. На практике используются также многослойные (MLP) — нейронные сети, а также KAN (Kolmogorov-Arnold) сети, состоящие из четырех и более слоев [Хайкин, 2006]:

$$\varphi(\boldsymbol{V},\boldsymbol{W},\boldsymbol{x}) = \sum_{l=1}^{L} v_l \left(\lambda\left(\sum \ ...\lambda\left(\sum_{m=1}^{M} w_{l,m} x_m\right)...\right)\right).$$

Чтобы обеспечить *универсальность аппроксиматора*, он должен обучаться в максимально широком классе сред при минимальной априорной информации о решении и свойствах среды.

2.3. Особенности АНС метода

В сравнении НС метода решения обратных с другими методами, например, методами тихоновской регуляризации и байесовским методом [Backus, 1988; Спичак, 2005], следует понимать, что последний так же, как и метод регуляризации, находит решение обратной задачи, соответствующее минимуму функционала невязки [Спичак, 2005; Жданов, 2007; Дмитриев, 2012]. Основными проблемами численного решения экстремальной задачи минимизации функционала невязки (равно как и задачи максимизации функции плотности апостериорной вероятности в байесовском методе) в общем нелинейном случае, как известно, являются: необходимость задания 1-го приближения (например, в байесовском методе необходимо задание априорной плотности распределения искомой характеристики среды), множественные локальные экстремумы и неединственность глобального экстремума целевого функционала и др. Строгое обоснование применяемых численных методов и построение формализованных алгоритмов инверсии возможно только на специальных подмножествах (например, на подмножестве монотонных решений и др.) [Тихонов и др., 1990]. В связи с этим для численного решения нелинейной экстремальной задачи, наряду с классическими методами оптимизации, на практике применятся эвристические методы, что часто приводит к субъективности получаемых результатов вследствие “человеческого фактора”. При этом “плохую экстремальную задачу” приходится решать каждый раз заново при изменении входных данных.

В НС методе приближенный обратный оператор строится заранее, а плохая экстремальная задача обучения НС аппроксиматора решается один раз для заданного класса сред; в качестве априорной информации о решении задается только детерминированный диапазон возможных значений искомой характеристики среды. После того как НС аппроксиматор обучен, он фиксируется и с его помощью решение обратной задачи может находиться формализованно и многократно для любых входных данных путем вычисления значений нейросетевого полинома по формуле (4), что составляет первые десятки секунд и по скорости кардинально превышает любой оптимизационный метод, предполагающий расчет множества прямых задач (онлайн) в процессе поиска экстремумов. Процедура нахождения решения обратной задача НС методом абсолютно детерминирована, формализована и лишена человеческого фактора. Для построения аппроксиматора в заданном классе сред нужна определенная квалификация пользователя в области НС технологий, при применении же обученного аппроксиматора к измеренным данным так же, как в традиционном палеточном методе, такая квалификация не требуется.

Достаточно большой объем суперкомпьютерного времени, требуемого для построения обучающего множества опорных решений, не очень важен (с методической точки зрения), так как эти операции производятся заранее, а производительность вычислительных кластеров со временем достаточно быстро возрастает.

Основная сложность реализации НС метода заключается в решении задачи обучения НС аппроксиматора с минимальной собственной ошибкой при сохранении детальности решения. НС метод не минимизирует невязку решения для конкретной правой части в явном виде (минимизируется ошибка решения обратной задачи на множестве опорных решений, т.е. ошибка обучения аппроксиматора). Если фактическая невязка решения при фиксированных измеренных данных оказывается “слишком” большой, то НС решение используется как 1-е приближение и уточняется каким-либо методом минимизации невязки, что требует дополнительных ресурсов по времени. В этом случае преимущество АНС метода заключается в быстром получении довольно точного (как показывает опыт) 1-го приближения.

2.4. Архитектура нейросети

В НС методах существенную роль играет архитектура и, в частности, число скрытых слоев нейронной сети, а также число нейронов в слоях. В упомянутых в начале введения и многих других работах было показано, что трехслойные нейросети с одним промежуточным слоем успешно применяются для решения малопараметрических обратных задач геоэлектрики с небольшим числом искомых параметров $N \le 15$. При этом в работах [Spichak, Popova, 2000; Спичак, Попова, 2005; Spichak, 2020] на примере обратной 3D-задачи геоэлектрики с числом искомых параметров $N = 6$ численно было показано, что добавление в нейросеть второго скрытого слоя практически эквивалентно соответствующему увеличению числа нейронов в скрытом слое трехслойной нейросети (для обучения нейросетей авторами строилось множество опорных решений, которое включало до 120 примеров различных комбинаций этих 6 параметров, которые изменялись на заданных сетках значений). Из теории известно, что, согласно теореме Цыбенко [Cybenko, 1989], любая непрерывная функция на единичном замкнутом кубе может быть аппроксимирована с любой наперед заданной точностью трехслойным персептроном с одним скрытым слоем (при неограниченном увеличении числа нейронов в промежуточном слое). Функция обратного оператора — отображение “вход–выход” — условно-корректной обратной задачи является непрерывной [Тихонов, Арсенин, 1979], поэтому данная теорема служит теоретическим обоснованием НС метода решения обратной задачи с использованием трехслойных нейросетей. Однако из указанной теоремы не следует, что один скрытый слой является оптимальным в смысле собственной ошибки обучаемого аппроксиматора, времени обучения, простоты реализации: для каждого типа нелинейности оптимальное число скрытых слоев определяется в индивидуальном порядке, при этом теорема существования аппроксимирующей нейросети остается в силе и для многослойных сетей [Хайкин, 2006; Ziming et al., 2024]. Вычислительный опыт авторов настоящей статьи показал, что при решении многопараметрических обратных задач геоэлектрики, в которых число искомых параметров составляет сотни и тысячи, трехслойные нейросети не являются достаточно эффективным инструментом, так как обучаются с большей ошибкой, чем, например, пятислойные. Уменьшить эту ошибку за счет увеличения числа нейронов в скрытом слое на практике не представляется возможным. Использование пятислойной сети с тремя скрытыми слоями может уменьшать ошибку обучения на 20% и более по сравнению с трехслойной (см. пример в п. 2.3). Дальнейшая модернизация НС метода решения многопараметрических обратных 2D-, 3D-задач с использованием пятислойной нейросети и специальных дополнительных преобразований данных, предшествующих самой задаче обучения нейросети, позволила решать обратные 2D-задачи геоэлектрики с числом искомых параметров $N \le 300$ [Шимелевич, Оборнев, 2009; Гужва и др., 2010; Шимелевич, 2013]. Объемы банков решений для таких задач составляли 10000 примеров различных вариантов сред. Результаты обобщения алгоритмов на обратные 3D-задачи геоэлектрики с числом искомых параметров $N \sim 1000$ и более представлены в работах [Шимелевич и др., 2017; 2018; 2022], при этом для обучения НС аппроксиматоров, строились банки опорных решений объемом до 25 000 примеров.

2.5. Рассматриваемые классы сред

Ясно, что с помощью кусочно-постоянных функций можно аппроксимировать любые распределения удельного сопротивления реальных сред с требуемой для практики точностью, что и предопределило рассмотрение обратной задачи в классах сред с кусочно-постоянным распределением удельного сопротивления. При такой аппроксимации конкретный геоэлектрический класс сред G_{S_N}, в котором решается обратная задача (1)–(2), определяется:

– размерами исследуемой области Ω, соответствующими типу решаемой геофизической задачи (малоглубинные, локальные, региональные задачи и др.);

– конфигурацией и размерностью N сетки параметризации θ_N, покрывающей область исследований Ω;

– диапазоном $[s_{\min}, s_{\min} + D_s]$ допустимых изменений логарифма удельного сопротивления $\lg\rho$ реальных сред в ячейках сетки θ_N в соответствии с (2); при наличии априорной информации диапазон изменений может задаваться индивидуально для отдельных групп параметров.

Элементами геоэлектрического класса G_{S_N} при фиксированных Ω, N, θ_N, $s_{\min}$, D_s являются модели сред с любой комбинацией удельных сопротивлений ρ^n, задаваемых в ячейках сетки θ_N при условии (2).

При задании геоэлектрического класса сред необходимо учитывать, что с ростом N при фиксированных размерах области Ω увеличивается детальность сетки параметризации, а значит, и детальность решений обратной задачи, однако при этом, как правило, растет неустойчивость решений (ухудшается обусловленность задачи). В то же время при малых N задача будет более устойчивой, но решения — слишком грубыми. Необходим компромисс между детальностью решения и его устойчивостью [Страхов, 1978; Дмитриев, 2012]. В работах [Шимелевич, 2013; 2020; Шимелевич и др., 2017; 2018] рассмотрены методы и примеры построения оптимальной (регуляризованной) сетки параметризации $\theta_{N\max}$, для которой достигается требуемый компромисс. Далее в данной работе обратная задача (1)–(2) рассматривается на регуляризованной сетке параметризации θ_{Nmax}.

Достаточно подробные обзоры по применению НС сетей в обратных задачах геоэлектрики представлены в работах [Raiche, 1991; Poulton, 2002; Спичак, Попова, 2005; Шимелевич, Оборнев 2007; 2009; 2018; Спичак, 2019; Spichak, 2020]. Обзор по применению АНС метода решения обратных задач других геофизических методов представлен в работе [Оборнев и др., 2020].

3. СХЕМА АЛГОРИТМА АНС МЕТОДА

Базовая схема одной из последних авторских версий алгоритма АНС метода решения обратной задачи геоэлектрики, представленная в работе [Шимелевич и др., 2018], включает следующие основные блоки:

I. Численное решение прямой задачи геоэлектрики.

II. Построение модельного класса сред на основе эффективной сеточной параметризации (адаптивная регуляризация задачи).

III. Построение (обучение) НС аппроксиматора.

IV. НС инверсия наблюденных данных; расчет невязки решения обратной задачи.

V. Построение уточняющих НС аппроксиматоров.

VI. Расчет апостериорных оценок степени неоднозначности финального решения.

В настоящей версии АНС алгоритма, представляемой в данной статье, изменены:

– III блок обучения НС аппроксиматора функционирует на основе авторской версии CNN.INV-сети сверточного типа;

– в блок V добавлен новый алгоритм уточнения НС решения методом случайного поиска с предварительной кластеризацией методом k-средних и применением интерполяционных процедур методом IDW [Шимелевич и др., 2022].

3.1. Структура нейросети CNN.INV сверточного типа

Классические MLP-сети являются полносвязными сетями: каждый элемент последующего слоя связан со всеми элементами предыдущего слоя, причем каждая связь имеет свой персональный весовой коэффициент. В современных нейросетевых технологиях применяются Глубокие нейронные сети (ГНС или DNN — Deep Neural Networks), состоящие из большого количества скрытых иерархических слоев с различными функциями нелинейных преобразований данных. Среди ГНС особое место занимают сверточные сети (CNN — Convolutional Neural Networks) [Bengio, 2009]. В сверточной нейронной сети каждый элемент последующего слоя данных связан с группой элементов (окном) предыдущего слоя посредством фильтра: окно, являясь “ядром” свертки, “скользит” по элементам (признакам) предыдущего слоя, а выход свертки определяет элементы последующего. На начальном этапе этот процесс выполняется непосредственно для первого слоя, то есть для анализируемого цифрового массива входных данных. После многократного прохождения указанных слоев сжатия данных и других специальных преобразований, привязанных к решаемой задаче, формируются *сжатые карты признаков*, которые содержат наиболее существенную информацию обо всем массиве исходных данных, что позволяет уменьшить ошибку обучения нейросети, а также облегчает и ускоряет ее обучение. Итоговые карты признаков передаются для обучения на вход обычной полносвязной MLP-сети.

В задачах геоэлектрики полная размерность входных данных, подаваемых на вход нейросети, для практических 2D-, 3D-задач может составлять $\sim 10^4 \div 10^6$, а размерность выходного слоя, определяемая числом искомых параметров модели $\sim 10^2 \div 10^4$. Обучить полносвязную MLP-сеть с такими размерностями входа и выхода не представляется возможным, поэтому сжатие данных и использование аппарата сверточных сетей необходимо.

В работах [Шимелевич, Оборнев, 2003; 2009] было введено понятие локального НС аппроксиматора при решении обратных коэффициентных задач геоэлектрики и задач геоэлектромониторинга, который устанавливает количественную связь некоторой выделенной группы (окна) входных данных с определенной группой (окном) параметров среды. Это позволяло решать так называемую задачу *частичной инверсии на основе неполных данных*. В дальнейших работах [Гужва и др., 2010; Шимелевич и др., 2017; 2018] с использованием техники локальных НС аппроксиматоров осуществлялось сжатие входных и выходных данных обучаемой нейросети. При этом окна параметров выделялись в пределах каждого i-го, $i = 1,...,I$, яруса сетки параметризации θ_N, покрывающей исследуемую область Ω. Кроме сжатия данных, при обучении аппроксиматоров в этих работах проводились дополнительные преобразования выделенных окон входных данных, в результате чего формировались поярусные *сжатые карты признаков*, которые подавались на вход обучаемых полносвязных пятислойных MLP-сетей, выходом которых является набор поярусных локальных НС аппроксиматоров. Совокупность проводимых преобразования данных, предшествующих обучению НС аппроксиматора, позволяет в разы уменьшить ошибку его обучения.

В данной статье представлена многослойная нейронная сеть CNN.INV (Convolutional Neural Networks.Inversion) сверточного типа для обучения НС аппроксиматоров применительно к решению нелинейных 3D обратных геоэлектрики. Аббревиатура CNN является общепринятой в литературе для обозначения сверточных сетей [Шарден и др., 2018], аббревиатура INV — подчеркивает прикладную специфику сети, ориентированную на обратные задачи геоэлектрики. Предлагаемая версия сети CNN.INV сверточного типа создана на базе предыдущих обучающих версий с некоторыми добавлениями, позволившими уменьшить ошибку обучения НС аппроксиматоров и повысить тем самым разрешающую способность и эффективность АНС метода для решения обратной задачи геоэлектрики.

На рис.1 представлена укрупненная схема сверточной CNN.INV-сети.

Замечание 1. *В НС технологиях последовательные преобразования данных традиционно называют слоями нейросети. В данной работе авторы придерживаются этой терминологии.*

На вход сверточной нейросети подаются массивы характеристик поля из рассчитанного заранее банка опорных решений $\mathbb{Q}_{bs}$; в процессе построения банка решений производится варьирование параметров задаваемой *фоновой* (горизонтально-слоистой) среды, окружающей исследуемую область Ω, для подавление ее влияния на обучаемый аппроксиматор.

Замечание 2. *Как известно* [Бердичевский, Дмитриев, 2009], *задание фоновой среды с простым строением, окружающей исследуемую аномальную область* Ω, *позволяет "избавиться" от бесконечных областей и сформулировать прямую краевую задачу в конечной области пространства. На практике реальная фоновая среда часто бывает неизвестной, в этом случае на основе априорной информации задается тип фоновой среды — некоторая гипотетическая упрощенная* (*обычно горизонтально-слоистая*) *модель среды, обеспечивающая постановку граничных условий и корректную разрешимость соответствующей краевой задачи. Варьирование неизвестных параметров фоновой среды в процессе обучения позволяет уменьшить ее влияние на свойства обучаемого аппроксиматора* [Шимелевич и др., 2018].

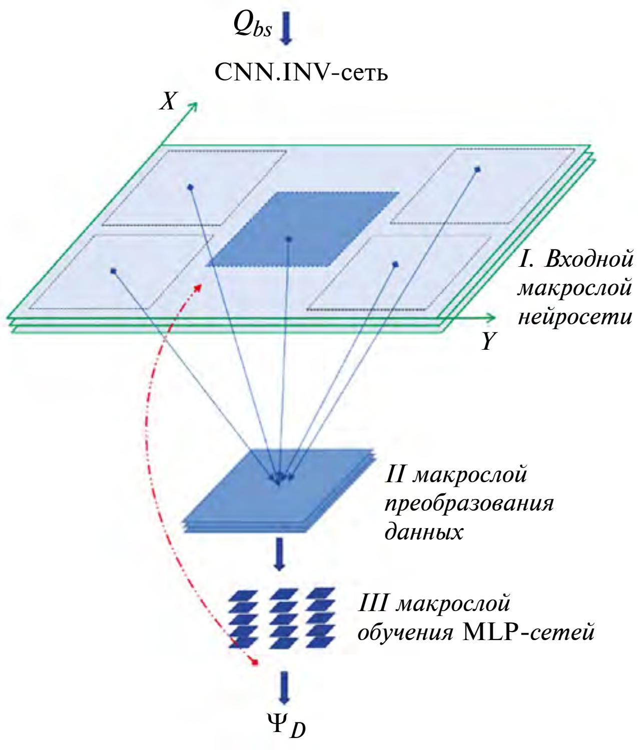

Рис. 1. Схема CNN.INV-сети сверточного типа для обучения НС аппроксиматора. Красная стрелка обозначает многократные циклы работы сети с целью выбора оптимальных значений параметров проводимых преобразований.

I. Входной макрослой нейросети представляет собой полный обучающий массив расчетных характеристик поля из банка опорных решений $\mathbb{Q}_{bs}$. Макрослой включает набор двумерных площадных массивов (на рис. 1 двумерные массивы схематически показаны зелеными прямоугольниками), каждый из которых состоит из величин компонент матриц импеданса и вектора типпера, заданных в области Ω_1 расположения площадной измерительной сети (по осям X, Y) на сетке частот. Число двумерных массивов, с учетом числа частот и числа используемых комплексных характеристик поля, может достигать 200 и более.

II. Макрослой преобразований данных включает следующие преобразования данных, направленные на уменьшение ошибки обучения НС аппроксиматора:

— Сжатие данных на основе выделения окон данных; в пределах каждого i-го, $i = 1,\dots,I$, яруса сетки параметризации θ_N выделяется группа параметров среды $s_i = (s_i^1,..,s_i^{N_i})$, образующая центральное окно параметров D_i (рис. 2); для этого окна параметров из всего массива входных данных выявляется группа $f_i = (f_i^1,..,f_i^{M_i})$ наиболее значимых (существенных) входных данных, которые образуют центральное сверточное окно входных данных F_i (рис. 2). В результате подобного сжатия размерность входных данных может быть уменьшена на несколько порядков. Оптимизируемыми параметрами преобразования сжатия в процесс работы сети являются размеры окон параметров D_i и входных данных F_i.

— Аугментация входных данных представляет собой процесс увеличения объема обучающей выборки на основе уже имеющихся данных. В данной сети это реализовано путем накопления данных в центральном окне данных F_i за счет добавления с весами данных из других окон, показанных пунктирными прямоугольниками на рис. 1, рис. 2. Аугментация данного типа (проводимая по всем ярусам сетки параметризации) позволяет увеличить объем обучающей выборки, оставаясь в рамках исходного банка решений, на 2–3 порядка в зависимости от па-

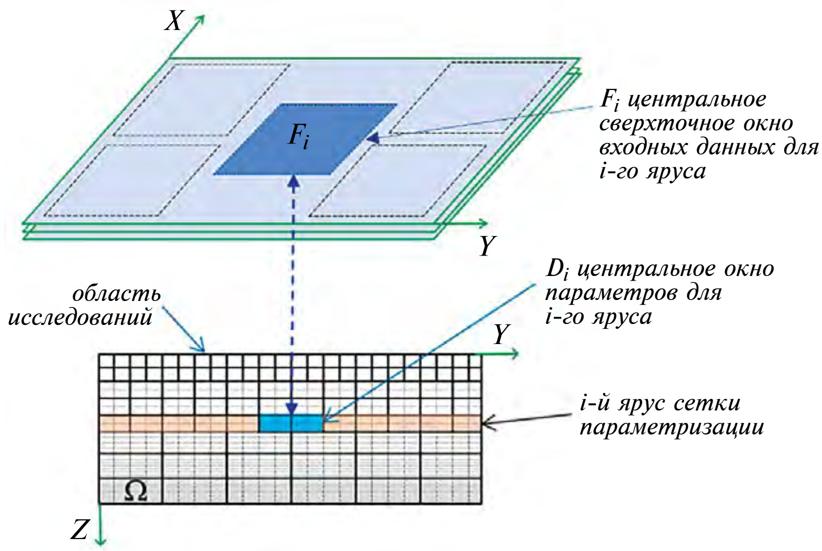

Рис. 2. Схема выделения окна параметров D_i и соответствующего окна данных F_i для i-го яруса сетки параметризации. В области исследований Ω толстыми линиями отмечена сетка параметризации θ_N, тонкими — сетка прямой задачи.

раметров выбираемых окон входных данных. Оптимизируемыми параметрами аугментации являются размеры окон входных данных F_i и итоговый объем обучающей выборки, получаемой в объединенном центральном окне.

– **Пулинг** заключается в дополнительном нелинейном сжатии входных данных для уменьшения размерности пространства обучающих данных и улучшения аппроксимационных свойств сети. В данной сети пулинг реализуется путем выделения квазиоднородных групп данных и объединении нескольких значений в одно по заданному правилу. Предусмотрены также иные известные правила пулинга [Шарден и др., 2018]. Оптимизируемыми критериями пулинга являются размерности квазиоднородных групп данных и правила объединения.

– **Построение локальных фокусирующих аппроксиматоров**, обучаемых на вспомогательных сетках параметризации, которые отличаются от исходной оптимальной регуляризованной сетки θ_N тем, что имеют увеличенную детальность для обучаемого i-го, яруса и укрупненную — для уже обученных ярусов с номерами $i-1$, $i-2$, ..., 1. Это позволяет увеличить общую детальность итогового решения как по вертикали, так и по горизонтали в заданной исследуемой области Ω; для каждого такого аппроксиматора рассчитывается дополнительный банк решений прямых задач. Оптимизируемыми параметрами данного преобразования является максимально возможная поярусная детальность решения при заданном пороговом ограничении на ошибку обучения.

– Выходом макрослоя преобразования данных являются поярусные сжатые карты признаков $\bar{F}_i$, которые подаются на вход обучаемых полносвязных MLP-сетей.

III. Макрослой обучения набора полносвязных MLP-сетей; входом каждой MLP-сети является карта признаков $\bar{F}_i$, выходом — центральное окно D_i параметров i-го, $i = 1,...,I$, яруса сетки параметризации θ_N; перед обучением сетей проводится нормализация входных данных (приведение измеренных данных к интервалу $[0,1]$).

Выбор оптимальной структуры нейросети [Шимелевич и др., 2018]. При решении задачи оптимизации структуры нейросети проявляется основное противоречие между сложностью сети и ее интерполяционными возможностями: простые сети с малым числом свободных параметров характеризуются большой интерполяционной ошибкой, сложные сети дают меньшую ошибку, но труднее обучаются. Критерием выбора наилучшей архитектуры нейросети служит величина собственной интерполяционной ошибки по экзаменационной выборке. Наилучшие результаты применительно к рассматриваемым задачам геоэлектрики показали пятислойные сети с количеством нейронов в скрытых слоях 32,16 и 8. В п. 2.3 приводится пример сравнения ошибок обучения трех- и пятислойных MLP-сетей.

Выходом III макрослоя и всей CNN.INV-сети является набор обученных локальных НС аппроксиматоров $\mathbf{\Psi}_D^i, i = 1, ..., I$ для каждого яруса сетки параметризации с рассчитанными собственными ошибками.

В процессе работы всей CNN.INV-сети реализуются многократные циклы (условно обозначенные красной стрелкой на рис. 1), при которых меняются числовые параметры всех проводимых преобразований (начиная от формирования окон данных), обучаются вспомогательные локальные аппроксиматоры и определяются их ошибки; в результате определяются оптимальные значения параметров преобразований, при которых собственная ошибка финального аппроксиматора минимальна.

Для обучения MLP-сетей используются авторская программа с участием авторов статьи [Гужва и др., 2010], а также бесплатные общедоступные библиотеки TensorFlow, Keras, PyTotch.

3.2. Решение обратной задачи с помощью локального аппроксиматора

Входные данные $\overline{F}_i$ локальных аппроксиматоров $\mathbf{\Psi}_D^i$ имеют пространственные размеры по осям X и Y обычно значительно меньшие, чем полная область Ω_1, на которой определены исходные входные данные при обучении аппроксиматоров. Применение к полным входным данным локальных поярусных аппроксиматоров $\mathbf{\Psi}_D^i, i = 1,...,I$, осуществляется путем перемещения входа $\overline{F}_i$ соответствующего i-го аппроксиматора по горизонтали в пределах области Ω_1 измеренных (модельных) данных таким образом, чтобы выход аппроксиматора в конечном итоге определил все параметры каждого i-го яруса. Так как при работе CNN.INV-сети производится максимальное подавление влияния среды, окружающей область Ω, на которой обучался локальный аппроксиматор, то в случае протяженных по горизонтали областей, аппроксиматор практически может применяться к любому участку $\Omega_{изм}$ измеренных данных, в том числе — выходящих за пределы области Ω_1 обучения аппроксиматора, и на своем выходе определять весь столбец D_i $i = 1,...,I$, окон параметров среды, расположенных под участком $\Omega_{изм}$. Данная техника инвертирования измеренных данных эффективно работает для неоднородных, но "более или менее" регулярных в горизонтальном направлении сред до глубин, определяемых нижней границей области обучения. Это позволяет решать обратные задачи геоэлектрики для измеряемых данных достаточно большой протяженности по горизонтали при условии, если, по имеющейся априорной информации, реальная среда не выходит за рамки класса сред, для которого обучался аппроксиматор.

4. ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ

4.1. Пример построения нейросетевого аппроксиматора на основе CNN.INV-сети

Обратная задача (1)–(2) рассматривается в классе сред G_1^{3D} с кусочно-постоянным распределением удельного сопротивления в ячейках заранее построенной регуляризованной сетки параметризации θ_N, покрывающей заданную исследуемую область Ω. Класс сред G_1^{3D} определяется сеткой параметризации θ_N, которая представлена на рис. 3, полная размерность сетки параметризации, определяющая общее число искомых параметров модели, составляет $N = 1332$. Для данного класса сред было принято, что логарифм удельного сопротивления в исследуемой области может меняться в пределах от 0 до 4, $0 \le \lg\rho \le 4$. В этом случае $s_{\min} = 0$, $D_s = 4$ и априорные ограничения (2) для параметров задачи принимают вид $[0 \le s^n \le 4]$. *Элементами класса G_1^{3D}, являются модели сред с любой комбинацией удельных сопротивлений ρ^n, $n = 1,...,N$ в ячейках указанной сетки параметризации θ_N при условии* $[0 \le \lg\rho^n \le 4]$.

Данный класс сред соответствует задачам МТЗ-АМТЗ малой и средней глубинности при поисках рудных и углеводородных месторождений [Сараев и др., 2010; Фельдман и др., 2008]. В качестве входных данных обратной задачи использовались четыре компоненты матрицы импеданса ($\boldsymbol{Z}_{xx}$, $\boldsymbol{Z}_{xy}$, $\boldsymbol{Z}_{yx}$, $\boldsymbol{Z}_{yy}$).

Для данного класса сред G_1^{3D} строился банк опорных решений. Параметры $s^n = \lg\rho^n$, $n = 1,...,N$ модели в каждой ячейке сетки параметризации варьировались независимо по равномерному закону в диапазоне $[0 \le s^n \le 4]$. Для каждой комбинации параметров модели решались прямые задачи и формировались соответствующие отклики поля при $z = 0$: действительные и мнимые части четырех компонент матрицы импеданса на сетке из 14 периодов в диапазоне от 0.001 до 21.54 с, на пространственной площадной сетке, включающей 900 точек (30×30). Общий объем банка опорных решений составил 25 000 примеров различных комбинаций параметров сред. Для решения прямой задачи использовался

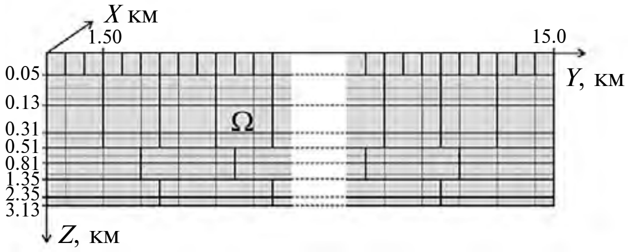

Рис. 3. Схема кусочно-постоянной параметризации класса сред G_1^{3D}. Тонкими линиями обозначена сетка прямой задачи, жирными — сетка параметризации.

модуль известной программы R. Mackie mtd3fwd, находящейся в открытом доступе.

В соответствии со схемой CNN.INV-сети для данного класса сред G_1^{3D} строились локальные НС аппроксиматоры. Средние по ярусу относительные ошибки обучения аппроксиматоров на экзаменационной выборке составили от 2.19 до 13.95%.

В качестве примера работы CNN.INV-сети в табл. 1 приведено сравнение ошибок обучения 5-слойной (32, 16, 8) и трехслойной (56) MLP-сетей (в скобках указано число нейронов в скрытых слоях) для одного и того же банка решений. Из табл. 1 видно, что относительное уменьшение ошибки обучения 5-слойной MLP-сети может достигать 52%. С другой стороны, увеличение числа нейронов в срытом слое трехслойной сети от 56 до 112 и 224 уменьшает, как правило, ошибку обучения лишь на десятые доли процентов, при этом время обучения возрастает на десятки процентов.

Технические параметры обучения сети на кластере:

1. Время расчета одной прямой задачи на всех используемых частотах: на персональном компьютере 20 мин, на суперкомпьютерном кластере 12 мин на одном вычислительном узле.

2. Расчет всего множества опорных решений из 25 000 примеров составил ~17 ч при использовании 300 ядер CPU, например, 6 пусков по 50 ядер.

3. Время обучения аппроксиматора на готовом множестве опорных решений составило ~72 ч на одном вычислительном узле без распараллеливания.

4.2. Результаты работы аппроксиматора

Модельные данные. С помощью набора обученных аппроксиматоров была решена обратная 3D-задача для синтезированных МТ данных, соответствующих модели высокоомного однородного полупространства (1000 Ом·м), содержащего низкоомное пирамидальное включение (1–10 Ом·м) (рис. 4а).

Результаты решения обратной задачи, представленные в табл. 2 и на рис. 4б, показывают, что вмещающая среда и пирамидальное включение, в том числе и точечные приповерхностные включения, определяются вполне удовлетворительно.

Обозначим через n — количество точек измерительной пространственной XY-сетки, Nt — число периодов, а $\tilde{e}_i(j,t)$ и $e_i'(j,t)$ $i=1,\ldots,n,\ t=1,\ldots,Nt$ — комплексные значения одной из четырех (индекс $j=1,\ldots,4$ обозначает одну из компонент тензора импеданса $\boldsymbol{Z}_{xx}$, $\boldsymbol{Z}_{xy}$, $\boldsymbol{Z}_{yx}$, $\boldsymbol{Z}_{yy}$) модельных и синтезированных электромагнитных характеристик соответственно. Для полученного решения вычислялась средневзвешенная относительная невязка:

$$\delta=\left(\sum_{j=1}^{4} w_j\right)^{-1}\sum_{j=1}^{4} w_j\left(\frac{1}{Nt}\sum_{t=1}^{Nt}\delta_{j,t}^2\right)^{1/2}, \qquad (5)$$

где $w_j=\sum_{t=1}^{Nt}\sqrt{\frac{1}{n}\sum_{i=1}^{n}|\tilde{e}_i(j,t)|^2},\ \ j=1,\ldots,4,$ — весовые коэффициенты для четырех электромагнитных характеристик, которые определяют значимость каждой из компонент тензора импеданса в соответствии с уровнем модельных характеристик по всему планшету измерений;

$$\delta_{j,t}^2=\frac{\sum_{i=1}^{n}|\tilde{e}_i(j,t)-e_i'(j,t)|^2}{\sum_{i=1}^{n}|\tilde{e}_i(j,t)|^2},\ j=1,\ldots,4,\ \ t=1,\ldots,Nt, —$$

квадраты относительных площадных невязок. Средневзвешенная относительная площадная

Таблица 1. Сравнение ошибок обучения пятислойной и трехслойной MLP-сетей на экзаменационной выборке

Ярус	2	3	4	5	6	7	8
Относительное уменьшение ошибки для 5-слойной сети, %	52.08	23.79	12.58	7.44	5.76	0.71	18.49

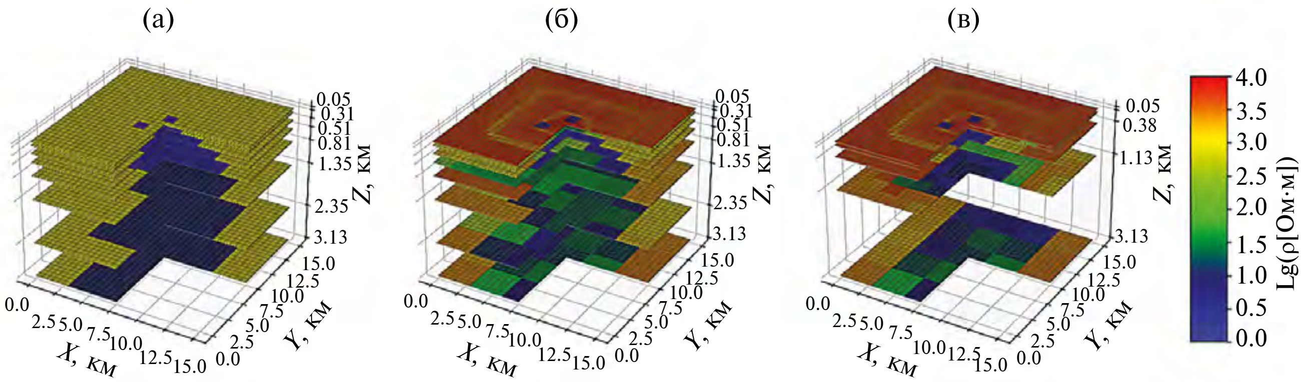

Рис. 4. Модель пирамидального высокопроводящего включения и точечных приповерхностных включений в плохо проводящем однородном полупространстве (а); результаты инверсии на 8-ярусной сетке параметризации (б); результаты инверсии на 5-ярусной сетке параметризации (в).

невязка (5) решения δ для данного модельного примера составила 22.32%. Относительно большая величина невязки означает, что разрешающая способность обученного НС аппроксиматора недостаточна для такой детальности решения; уточнение решения в данном случае не проводилось.

На рис. 4в приведены результаты НС инверсии, полученные для этой модели на основе тех же самых модельных данных с помощью предыдущей авторской версии нейросети, обученной ранее на основе 5-ярусной сетки параметризации, покрывающей такую же область исследований Ω [Шимелевич и др., 2022]; в этом случае средняя относительная невязка решения составила 50.44%. Столь большая относительная невязка решения 50.44% при использовании предыдущей версии обучения НС аппроксиматора для данной модели среды обусловлена более грубой аппроксимацией модели среды с использованием пятислойной сетки параметризации.

По сравнению с предшествующим результатом [Шимелевич и др., 2022], существенно увеличена детальность решения при меньших невязке и ошибке решения: добавлены три яруса сетки параметризации в интервале глубин от ~ 0.3 от до 3.13 км, которых не было в старой модели. Детальность решения обратной задачи по числу искомых параметров на глубинах от 50 до 3130 м для 8-слойной модели увеличилась по сравнению с 5-слойной моделью в 1.3 раза. Достичь такой детальности решений с помощью предыдущей версии нейросети не представлялось возможным: нейросеть не обучалась с приемлемой ошибкой. Сравнение результатов свидетельствует о существенно более высокой разрешающей способности CNN.INV-сети по глубине исследований, по сравнению с предыдущей версией сети.

Полевые данные. В качестве инвертируемых площадных натурных данных были взяты наблюдения Северо-Пясинской площади, которые являются частью региональной съемки Енисей-Хатангского прогиба. Размер области измерений 9.5×10 км. Съемка содержит 10 профилей, расстояние между профилями и пикетами примерно 1 км. Эти данные, предоставленные Центром ЭМИ, прошли первичную обработку, заключающуюся в фильтрации и нормализации [Фельдман и др., 2008].

После получения первого нейросетевого приближения на 8-ярусной сетке параметризации было проделано 5 итераций методом случайного поиска [Шимелевич и др., 2022]. На рис. 5 показано уменьшение относительной невязки по итерациям. На каждой итерации объем новой дополнительной выборки случайного поиска составил около 2000 решений прямых задач.

Таблица 2. Средние по ярусам относительные ошибки логарифма удельного сопротивления, полученные с помощью АНС метода, для модельного примера

№ яруса сетки параметризации	1	2	3	4	5	6	7	8	Среднее
Ошибка, %	15.65	4.71	3.80	12.79	13.93	13.95	12.14	10.24	10.90

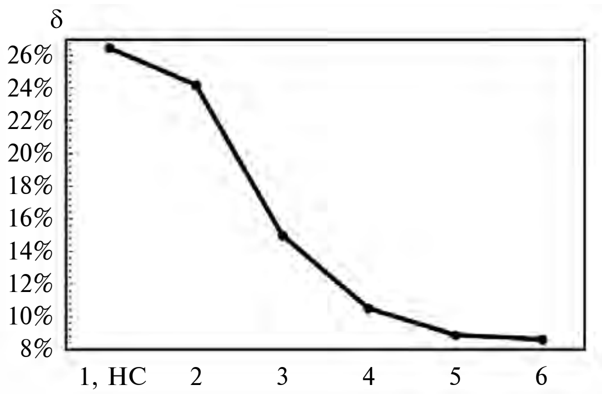

Рис. 5. График уменьшения средневзвешенных относительных невязок приближенного решения, уточняемого методом случайного поиска по итерациям.

На первой итерации (первичное НС решение) средняя относительная площадная невязка δ составила 26.2%, а на последней шестой итерации — 8.61%. На рис. 6 представлены горизонтальные сечения решения, полученного на 6-й итерации. Цветами обозначена шкала десятичного логарифма удельного сопротивления. Для сравнения, при НС инверсии на базе предыдущей версии для 5-ярусной сетки параметризации для тех же полевых данных на последней итерации была достигнута средняя относительная невязка 8.84%. Таким образом, с помощью построенной CNN.INV-сети удалось построить аппроксиматоры, позволяющие получить существенно более детальное решение обратной задачи при несколько меньшей невязке.

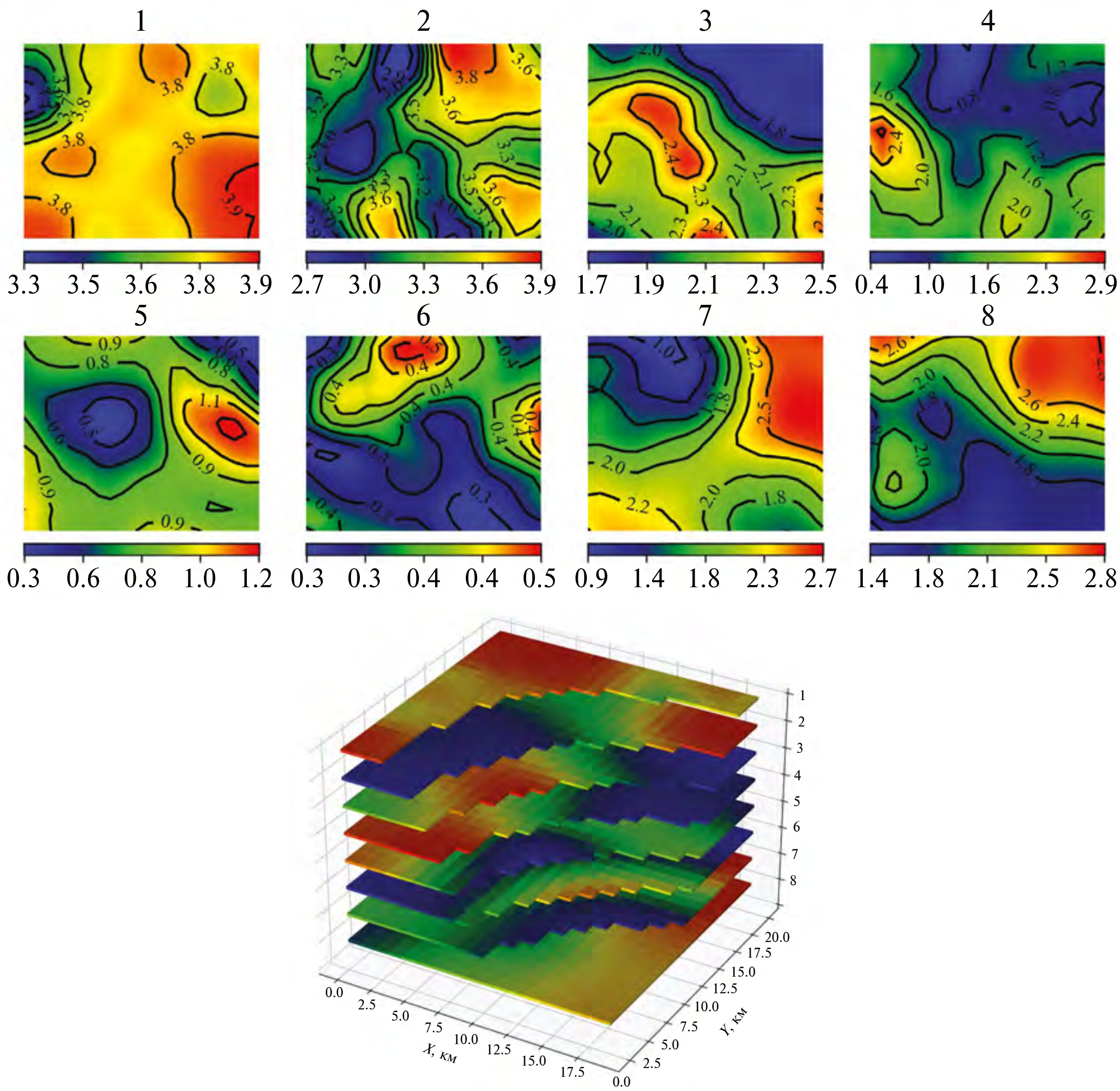

Рис. 6. Инверсия натурных данных Северо-Пясинской площади. Цифрами 1–8 обозначены поярусные горизонтальные сечения решения на шестой итерации.

5. ВЫВОДЫ

1. Представлены элементы алгоритма современного регуляризованного АНС метода решения нелинейной 3D обратной задачи геоэлектрики, связанные с построением и обучением авторской нейронной сети сверточного типа, адаптированной к свойствам решаемой задачи.

2. Приведен пример обучения нейросети для решения обратной 3D-задачи геоэлектрики в классе сред, соответствующем задачам МТЗ-АМТЗ средней глубинности ~3–4км при поисках рудных и углеводородных месторождений. Работа АНС метода продемонстрирована на модельных и полевых данных.

3. Полученные результаты позволяют сделать вывод о применимости современной версии НС метода для решения практических обратных 3D задач геоэлектрики с числом искомых параметров ~1000 и более с приемлемой для практики невязкой и детальностью в режиме реального времени, без задания 1-го приближения.

БЛАГОДАРНОСТИ

Работа выполнена с использованием вычислительных ресурсов Межведомственного суперкомпьютерного центра Российской академии наук (МСЦ РАН) и с использованием оборудования Центра коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ имени М.В. Ломоносова.

ФИНАНСИРОВАНИЕ РАБОТЫ

Исследование выполнено за счет гранта Российского научного фонда № 19-11-00333, https://rscf.ru/project/19-11-00333/.

СПИСОК ЛИТЕРАТУРЫ

Бердичевский М.Н., Дмитриев В.И. Модели и методы магнитотеллурики. М.: Научный мир. 2009. 680 с.

Гужва А.Г., Доленко С.А., Оборнев Е.А. и др. Использование адаптивных алгоритмов отбора существенных признаков при нейросетевом решении обратной задачи электроразведки // Нейрокомпьютеры: разработка, применение. 2010. № 3. С. 46–54.

Дмитриев В.И. Обратные задачи геофизики. Монография. М.: МАКС Пресс. 2012. 340 с.

Жданов М.С. Теория обратных задач и регуляризации в геофизике. М.: Научный мир. 2007. 712 с.

Корженевский А.В. Использование искусственных нейронных сетей для решения обратных задач электроимпедансной и магнитоиндукционной томографии // Журнал радиоэлектроники. 2001. № 12. [Электронный ресурс http://jre.cplire.ru/jre/dec01/index_e.html]

Оборнев Е.А., Оборнев И.Е., Родионов Е.А., Шимелевич М.И. Применение нейронных сетей в нелинейных обратных задачах геофизики // Журнал вычислительной математики и математической физики. 2020. Т. 60. № 6. С. 1053–1065. DOI: 10.31857/S0044466920060071

Оборнев Е.А., Шимелевич М.И., Доленко С.А., Шугай Ю.С. Классификация магнитотеллурических данных с использованием нейросетевого метода // Изв. высших учебных заведений. Геология и разведка. 2007. № 5. С. 60–68.

Ротштейн А.П., Митюшкин Ю.И. Идентификация нелинейных зависимостей нейронными сетями // Проблемы бионики. 1998. №49. С.168–174.

Сараев А.К., Антащук К.М., Никифоров А.Б., Прокофьев О.М., Романова Н.Е., Денисов Р.В. Возможности аудиомагнитотеллурических зондирований при решении структурных и поисковых задач в алмазоносных районах // Вопросы геофизики. 2010. Т. 443(43). С. 126–158.

Спичак В.В., Борисова В.П., Фрейнберг Э.Б. и др. Трехмерная электромагнитная томография эльбрусского вулканического центра по магнитотелургическим и спутниковым данным //Вулканология и сейсмология. 2007. № 1. С. 58–73.

Спичак В.В. Электромагнитная томография земных недр. М.: Научный мир. 2019. 374 с.

Спичак В.В. Трехмерная байесовская инверсия. Электромагнитные исследования земных недр / Спичак В.В. (ред.). М.: Научный мир. 2005. С. 91–109.

Спичак В.В., Попова И.В. Методология нейросетевой инверсии геофизических данных // Физика Земли. 2005. № 3. С. 71–85.

Спичак В.В., Попова И.В. Применение нейросетевого подхода для реконструкции параметров трехмерной геоэлектрической структуры // Физика Земли. 1998. № 1. С. 39–45.

Тихонов А.Н., Гласко В.Б., Дмитриев В.И. Математические методы в разведке полезных ископаемых. М.: изд-во "Знание". Сер. Математика и кибернетика. 1983. № 12. 64 с.

Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г. Численные методы решения некорректных задач. М.: Наука. 1990. 232 с.

Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука. 1979. 286 с.

Фельдман И.С., Окулесский Б.А., Сулейманов А.К. и др. Электроразведка методом МТЗ в комплексе региональных нефтегазопоисковых работ в европейской части России // Записки Горного института. 2008. Т. 176. С. 125–131.

Хайкин С. Нейронные сети: полный курс, 2-е издание. Пер. с англ. М.: ООО "И.Д. Вильямс". 2006. 1104 с.

Цымблер М.Л., Юртин А.А. Восстановление пропущенных значений временного ряда на основе совместного применения аналитических алгоритмов и нейронных сетей // Вычислительные методы и программирование. 2023. Т. 24. № 3. С. 243–259. DOI:10.26089/NumMet.v24r318

Шарден Б., *Массарон Л.*, *Боскетти А.* Крупномасштабное машинное обучение вместе с Python. М.: ДМК Пресс. 2018. 358 с.

Шимелевич М.И., *Оборнев Е.А.*, *Оборнев И.Е.*, *Родионов Е.А.* Аппроксимационный нейросетевой метод решения многомерных нелинейных обратных задач геофизики // Физика Земли. 2017. № 4. С. 100–109.

Шимелевич М.И. Методы повышения устойчивости инверсии данных геоэлектрики на основе нейросетевого моделирования // Геофизика. 2013. № 4. С. 49–55.

Шимелевич М.И. О методе расчета модуля непрерывности обратного оператора и его модификаций с приложением к нелинейным задачам геоэлектрики // Вычислительные методы и программирование. 2020. Т. 21. Вып. 4. С. 350–372. DOI: 10.26089/NUMMET.V21R430

Шимелевич М.И., *Оборнев Е.А.* Аппроксимационный метод решения обратной задачи МТЗ с использованием нейронных сетей // Физика Земли. 2009. № 12. С. 22–38.

Шимелевич М.И., *Оборнев Е.А.* Аппроксимационный подход к решению обратной задачи геоэлектрики с использованием нейронных сетей. Тез. докл. Международная конференция “Обратные и некорректные задачи математической физики”. Новосибирск, Академгородок, 5–12 августа, 2012. С. 289.

Шимелевич М.И., *Оборнев Е.А.* Нейросетевая инверсия МТ данных в классах параметризованных геоэлектрических разрезов // Физика Земли. 2007. № 3. С. 25–30.

Шимелевич М.И., *Оборнев Е.А.* Применение метода нейронных сетей для аппроксимации обратных операторов в задачах электромагнитных зондирований // Известия высших учебных заведений. Геология и разведка. 1999. № 2. С. 102–106.

Шимелевич М.И., *Оборнев Е.А.*, *Оборнев И.Е.*, *Родионов Е.А.* Алгоритм решения обратной задачи геоэлектрики на основе нейросетевой аппроксимации // Сибирский журнал вычислительной математики. 2018. Т. 21. № 4. С. 451–468. DOI: 10.15372/SJNM20180408

Шимелевич М.И., *Родионов Е.А.*, *Оборнев И.Е.*, *Оборнев Е.А.* Нейросетевая 3D инверсия полевых данных геоэлектрики с расчетом апостериорных оценок // Физика Земли. 2022. № 5. С. 3–13. DOI: 10.31857/S0002333722050246

Backus G. E. Bayesian inference in geomagnetism // Geophys. J. Int. 1988. V. 92(1). P. 125–142.

Bengio Y. Learning Deep Architectures for AI // Foundations and Trends in Machine Learning. V. 2. Is. 1. P. 1–127. 2009. DOI:10.1561/2200000006

Cybenko G. Approximation by superpositions of a sigmoidal function // Math. Control Signals Syst. 1989. V. 2 (4). P. 303–314.

Hidalgo H., *Gomez–Trevino E.*, *Swiniarski R.* Neural network approximation of an inverse functional. Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN’94). 1994. V. 5. P. 3387–3392.

Obornev I.E., *Shimelevich M.I.*, *Obornev E.A.*, *Dolenko S.A.*, *Rodionov E.A.* Method of focusing neural network approximators for solving a nonlinear inverse problem of geophysics. Conference Proceedings, 1st Regional Conference on Data Science in Oil and Gas 2020. P. 1–6. DOI: 10.3997/2214-4609.202054022

Poulton M., *Glass C.*, *Sternberg B.* Recognizing EM ellipticity patterns with neural networks. 59th Ann. Internat. Mtg., Soc. Expl. Geophys. 1989. P. 208–212.

Poulton M., *Sternberg B.*, *Glass C.* Neural network pattern recognition of subsurface EM images // Journal of Applied Geophysics. 1992. V. 29. Is. 1. P. 21–36.

Poulton M.M. Neural networks as an intelligence amplification tool: A review of applications // Geophysics. 2002. V. 67. Is. 3. P. 979–993. DOI: 10.1190/1.1484539

Raiche A. A pattern recognition approach to geophysical inversion using neural nets // Geophysics J. Int. 1991. V. 105. Is. 3. P. 629–648.

Shimelevich M., *Obornev E.* The method of neuron network in inverse problems MTZ. Abstracts of the 14-th workshop on Electromagnetic Induction in the Earth, Sinaia. Romania, 1998. P. 159.

Ziming Liu, *Yixuan Wang*, *Sachin Vaidya*, *Fabian Ruehle*, *James Halverson*, *Marin Soljačić*, *Thomas Y. Hou*, *Max Tegmark*. KAN: Kolmogorov-Arnold Networks. arXiv. 2024. DOI: 10.48550/arXiv.2404.19756

Spichak V., *Popova I.V.* Artificial neural network inversion of MT-data in terms of 3D earth macro-parameters // Geoph. J. Int. 2000. V. 142. P. 15-26.

Spichak V.V. Computational Geo-Electromagnetics: Methods, Models and Forecasts. Cambridge, MA, USA: Elsevier. 2020. 462 p.

Application of Convolutional Neural Networks in Inverse Problems of Geoelectrics

M. I. Shimelevich[a, *], E. A. Rodionov[a, c, **], I. E. Obornev[a, b, *], and E. A. Obornev[a, ****]**

[a]Sergo Ordzhonikidze Russian State University for Geological Prospecting, Moscow, 117997 Russia
[b]Skobeltsyn Institute of Nuclear Physics, Moscow State University, Moscow, 119991 Russia
[c]Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Moscow, 123242 Russia
**e-mail: shimelevich-m@yandex.ru*
***e-mail: evgeny_980@list.ru*
****e-mail: o_ivano@mail.ru*
*****e-mail: obornevea@mail.ru*

Received November 25, 2023; revised June 24, 2024; accepted June 25, 2024

Abstract — Neural networks (NNs) are successfully used to solve inverse and other problems in geophysics. The aim of this work, which is a continuation of a series of works by a group of authors, is to improve the efficiency of the NN method for solving nonlinear inverse 3D problems of geoelectrics, based on the construction of the author's convolutional neural network. The network includes a number of additional special transformations (data compression, suppression of the influence of an unknown background environment, etc.) preceding the training of a classical MLP neural network and adapted to the inverse problem that is being solved. This allows us to formally, excluding the human factor, solve inverse problems of geoelectrics of large dimensions without specifying a first approximation based on data measured in areas whose dimensions exceed the dimensions of the network training area. The inversion speed is a few tens of seconds and does not depend on the physical dimensionality (2D or 3D) of the data. The solution to the inverse problem found using a trained neural network can, if necessary, be refined using a random search method. Numerical results of solving 3D geoelectric problems on model and field data are presented, confirming the stated development parameters.

Keywords: MTZ, 3D inverse problem, geoelectrics, neural networks, convolutional networks, NN approximator, refinement of neural network solution