

Раздел IV. Машинное обучение и нейронные сети

УДК 004.032.26

DOI 10.18522/2311-3103-2026-1-235-249

Б.В. Костров, С.И. Бабаев, А.И. Ефимов, В.Ю. Тарасова

РАЗРАБОТКА ГИБРИДНОЙ АРХИТЕКТУРЫ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ С ИСПОЛЬЗОВАНИЕМ СПЕКТРАЛЬНЫХ ПРЕОБРАЗОВАНИЙ

Предлагается архитектура гибридной свёрточной нейронной сети, сочетающей спектральные и пространственные слои, а также новый метод подвыборки (WalsPooling). Разработанная система применяется для оценки географической близости пары изображений на основе их визуального сходства. Пара изображений получена с разных сенсоров и определение их визуального сходства осложняется условиями съёмки с разным масштабом, углом наклона сенсора. На основе фрагмента съёмки на малой высоте выполняется поиск в базе данных изображений подстилающей поверхности. Поиск осуществляется в окрестности заданного маршрута на основе вектора признаков изображений, который формируется на последнем слое свёрточной нейронной сети. Система использует сиамскую архитектуру, так как на вход необходимо подавать пару изображений. Актуальность задачи обусловлена необходимостью обеспечения навигации БПЛА в условиях отсутствия или ненадёжности GPS-сигнала. Также рассматривается подход к формированию набора данных и его предварительная обработка. Поиск в базе данных выполняется в окрестности маршрута, что сокращает вычислительные затраты. Эксперименты включают анализ применимости предложенных слоёв (WalsPooling) и сравнение с традиционными методами подвыборки. Также в работе представлен метод линейной аппроксимации с обучаемыми параметрами для уменьшения размерности свёрточного слоя. Основное преимущество подхода – устойчивость к изменениям масштаба и угла съёмки за счёт комбинации спектральных и пространственных признаков. Результаты демонстрируют применимость метода для навигации БПЛА в условиях потери или ненадёжности GPS-сигнала. В результате эксперимента было получено, что использование восстановленных после спектрального преобразования изображений позволяет достичь наилучших результатов сходимости нейронной сети и среднеквадратической ошибки. Разработанная архитектура демонстрирует устойчивость к геометрическим и яркостным искажениям, а ее метрики качества (Precision = 0.728, Recall = 0.800, F1 = 0.872) подтверждает эффективность подхода для задач визуальной локализации на основе изображений из базы данных подстилающей поверхности.

Гибридная СНС; преобразование Уолиша; визуальная локализация; сиамские сети; визуальный поиск; WalsPooling.

B.V. Kostrov, S.I. Babaev, A.I. Efimov, V.Yu. Tarasova

CONVOLUTIONAL NEURAL NETWORK HYBRID ARCHITECTURE DEVELOPMENT USING SPECTRAL TRANSFORMATIONS

The hybrid convolutional neural network architecture with combining spectral and spatial layers, as well as new methods of subsampling (WalsPooling) and convolution (ConvWals) are proposed. The developed system is used to geographical proximity assess of images pair based on their visual similarity. A pair of different sensors obtained images visual similarity determination is complicated by different scales and sensor tilt angles shooting conditions. Based on the low-altitude image fragment, a search in the database of underlying surface images is performed. The search is performed in the surrounding area of a given route based on the vector of image features, which is formed on the last layer of the convolutional neural network. The system uses the Siamese architecture, since a pair of images must be submitted to the input. The relevance of this problem stems from the need to ensure UAV navigation in the absence or unreliability of a GPS signal. The approach to data set formation and its preprocessing is also considered. The database search is

performed in the surrounding area of the route, which reduces computational costs. The experiments include an analysis of the applicability of the proposed layers (WalsPooling, ConvWals) and a comparison with traditional pooling and convolution methods. The paper also presents a linear approximation method with trainable parameters for reducing the dimensionality of the convolutional layer. The main advantage of the approach is its resistance to changes in the scale and angle of shooting due to a combination of spectral and spatial features. The results demonstrate the applicability of the method for UAV navigation in conditions of loss of GPS signal is lost or unreliable. The experiment demonstrated that using images reconstructed after spectral transformation yields the best neural network convergence and mean square error. The developed architecture demonstrates robustness to geometric and brightness distortions, and its quality metrics (Precision = 0.728, Recall = 0.800, F1 = 0.872) confirm the effectiveness of the approach for visual localization tasks based on images from a surface database.

Hybrid SNS; Walsh transformation; visual localization; Siamese networks; visual search; WalsPooling.

Введение. Данная работа посвящена разработке архитектуры гибридной свёрточной нейронной сети (СНС) с новыми типами слоёв и методами обработки изображений. В работе предлагается новая структура гибридной СНС (сочетание спектральных и стандартных слоёв). Стандартные слои СНС применяются для выделения локальных признаков, спектральные – для глобальных. Предложен новый метод подвыборки (пулинга) на основе линейной аппроксимации с обучаемым параметром (модель сама выбирает к какой точке аппроксимировать), который в сочетании с WalsPooling улучшает сохранение пространственно-спектральных признаков.

В данной работе описывается способ взаимодействия компонентов (сиамская сеть, среднеквадратическое отклонение), а также предварительная обработка данных, которая состоит в применении спектрального преобразование Уолша, удалении низкочастотной составляющей. Предлагается использовать сиамскую архитектуру, так как необходимо только по визуальному сходству пары изображений оценить степень их географической близости. Степень близости оценивается как среднеквадратическое отклонение векторов признаков.

Данная архитектура применяется в задаче локализации БПЛА в условиях отсутствия GPS-сигнала на основе визуального сходства пар аэрофотоснимков. На вход нейронной сети подаются два изображения: текущий снимок с БПЛА (получен в реальном времени во время полёта) и спутниковый снимок из базы данных подстилающей поверхности (геопривязанный). Известны маршрут полета (стартовая и конечная точки) и GPS-координаты в начальный момент, но в процессе полёта сигнал GPS недоступен или ненадёжен. Цель работы оценить степень географической близости двух снимков на основе их визуального сходства, чтобы определить текущее положение БПЛА и при необходимости скорректировать траекторию. В данной работе рассматривается только задача сравнения изображений (оценка визуального сходства), а не полная навигационная система.

Обзор существующих решений. В работах [1–4] предлагается заменить классические методы подвыборки (операции уменьшения размерности) для сохранения большего объёма информации. В работе [1] применяется интерполяционный пулинг, учитывающий веса ближайших пикселей, что близко к предложенной линейной аппроксимации, однако кроме учета веса ближайших пикселей, также рассматривается обучаемый параметр для смещений (помимо весов еще выбирается к какому пикселю смещаться при аппроксимации). Предложенная линейная интерполяция с обучаемым параметром, так же как и в работах [1, 3, 4] направлена на минимизацию потерь важных признаков при уменьшении размерности. В работе [2] акцент сделан на ускорение вычислений. В работах [3, 4] используется вейвлет-пулинг, который также сохраняет структурную информацию за счёт спектрального разложения. Ключевым отличием предложенного WalsPooling, является то, что он применяется совместно с линейной аппроксимацией, что уникально и позволяет добиться сохранения пространственно-спектральных признаков. Вейвлет-пулинг [3, 4] использует другое спектральное представление, и в отличие от спектрального преобразования Уолша, он требует больше ресурсов и не адаптирован для сравнения изображений. Работы [1–4] тестируются на задачах классификации, с использованием набора данных

ImageNet [5], а не парном сравнении как в данном исследовании. Линейная аппроксимация имеет сходство с интерполяционным пулингом [1], но в данной работе смещение предлагается сделать обучаемым и для каждой соседней точки задать обучаемые веса (взвешенное суммирование). Предложенный метод имеет высокую устойчивость к искажениям, в отличие от рассмотренных работ.

В работах [6–8] применяются вейвлет-преобразования для уменьшения потери информации в СНС. В [6] вейвлеты встроены в нейронную сеть U-Net для задач восстановления изображений. В [7] вейвлеты используются для предобработки и ускорения СНС. В работе [8] используется вейвлет-преобразование Хаара для шумоподавления и выделения признаков. В работе [9] для анализа одномерных сигналов в нейросетях применяется ряд Фурье.

В работах [6–8] используются спектральные преобразования для улучшения СНС. Данные работы направлены на сохранение информации при уменьшении размерности. Однако, преобразование Уолша менее ресурсоемкое, чем вейвлет-преобразование и преобразование Фурье, и подходит для обработки изображений с БПЛА. Гибридная архитектура адаптирована под сравнение пар изображений, а не под классификацию/восстановление.

В работе [10] исследуется вейвлет анализ для обработки изображений с БПЛА, включая задачи навигации. Авторы комбинируют СНС с частотным разложением для улучшения выделения признаков. В [10] применяется бинарная классификация (обнаружение объектов). В отличие от [10], где вейвлеты применяются для предобработки, наш метод интегрирует преобразование Уолша непосредственно в обучаемую архитектуру, что повышает адаптивность к вариациям входных данных.

В работе [11] решается задача сопоставления изображений, снятых в схожих условиях. Предложенная архитектура нейронной сети работает с парой изображений, полученных с разных сенсоров, что критически важно для автономной навигации в условиях отсутствия GPS. Работа [12] посвящена обзорам проблем навигации, которые необходимо учитывать в задачах автонавигации. В работе [13] рассматривается двухуровневая система управления. В исследованиях [11, 14] используют стандартную архитектуру и классическими методами подвыборки.

Визуальный поиск. Определение визуального сходства – это задача, которая выявляет насколько идентичны два изображения, используя их визуальное содержание. Дано m изображений из поисковой базы данных $G = \{I_1, I_2, \dots, I_M\}$. Каждое изображение $I_j, j = \overline{1, M}$ характеризуется вектором числовых значений признаков $\Phi(I) = \langle \varphi_1(I), \varphi_2(I), \dots, \varphi_N(I) \rangle$, где N – число признаков (мощность пространства признаков) R^N . Определим $\tau_{k,s} = \tau(I_k, I_s)$ как показатель сходства, который измеряет насколько подобны два изображения $I_k \in G$ и $I_s \in G$. Решаемая задача состоит в изучении функций визуального сходства τ , которые вычисляют значение показателя сходства, основываясь на векторах числовых значений признаков, как:

$$\tau_{k,s} \geq \tau_{min}, \quad (1)$$

где τ_{min} – пороговое значение сходства.

Решение задачи определения визуального сходства состоит из следующих этапов:

1. Формирование вектора числовых значений признаков изображения, которые определяют наиболее заметную визуальную информацию;
2. Численная оценка визуального сходства (мера близости) на основе сформированных векторов [15, 16].

Для реализации указанных этапов применим сиамскую архитектуру нейронной сети. Так как в качестве входных значений необходимо подавать пару изображений, то разработанную архитектуру расширим до сиамской. Выходным значением, является численная оценка сходства, чем она меньше, тем изображения визуально ближе. В нашем случае рассматривается географическая близость. Географическая близость двух изображений, находящихся в одной локации, но полученные с разных сенсоров должны $\rightarrow 0$.

Сиамская архитектура. Сиамская архитектура нейронной сети – это особый тип нейронных сетей, который предназначен для сравнение пары входных данных и определения степени их схожести. Данная архитектура применяется в задачах, где необходимо измерить сходство или расстояние между объектами [17, 18].

К ключевым особенностям данной архитектуры можно отнести:

1. Две идентичные подсети («близнецы»):
 - ◆ обе ветви используют одинаковые веса;
 - ◆ обучаются параллельно на парах входных данных.
2. Функции сравнения:
 - ◆ контрастная (Contrastive Loss);
 - ◆ triplet Loss;
 - ◆ косинусное расстояние:

$$\text{cosine}_{\text{similar}}(A, B) = \frac{(A, B)}{\|A\| \cdot \|B\|}, \quad (2)$$

где (A, B) – скалярное произведение векторов A и B , $\|A\|$ – норма (длина) вектора A , вычисляемая как $\sqrt{\sum_{i=1}^m A_i^2}$, i – индекс суммирования, M – размер каждого из векторов A и B , $\|B\|$ – норма (длина) вектора B , вычисляемая аналогично.

- ◆ MSE (Mean Squared Error) – среднеквадратичная ошибка.

3. Выходное значение СНС – степень схожести: число от 0 до 1, где 0 – это идентичные объекты.

На рис. 1 приведена модель вычисления сходства с использованием сиамских сетей [19].

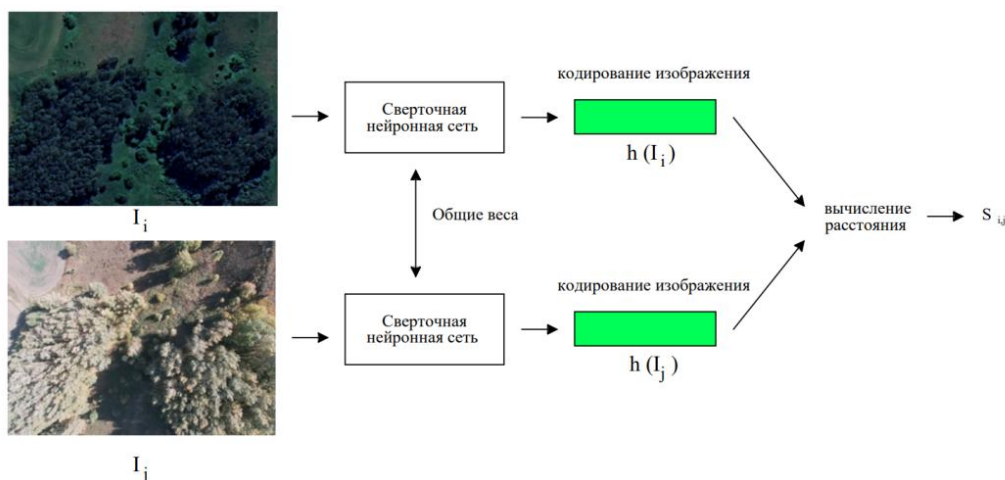


Рис. 1. Модель вычисления сходства с использованием сиамских сетей

Операции для уменьшения размерности. В данной работе рассматривались следующие операции для уменьшения размерности:

1. **Линейная интерполяция** – альтернативный подход к подвыборке, при котором используется взвешенная линейная интерполяция значений к некоторой точке. Модель на основе обучения выбирает к какой точке ближе аппроксимировать.

Алгоритм работы линейной интерполяции с обучаемым параметром.

1. Разделить входной тензор на блоки 2×2 .
2. Вычислить взвешенные коэффициенты для каждого пикселя в блоке, чтобы приблизить значение, например, к точке x_{11} .
3. Задаем смещение, как обучаемый параметр, что позволяет модели автоматически выбирать, к какой точке внутри блока аппроксимировать.

4. Для каждой точки в блоке задаем обучаемые веса, которые определяют их вклад в распределении.
5. Сформировать выходной тензор меньшего размера.
6. Передать выходной тензор дальше по нейронной сети.

Данный подход уменьшения размерности является альтернативой AvgPooling, но применяется тогда, когда важна именно плавность изменений, например, при спектральных преобразованиях. Данный подход может быть применен к нейронным сетям с малым количеством данных, так как в них меньше риск к переобучению. К преимуществам использования можно отнести то, что данная операция меньше искажает границы объектов, чем AvgPooling. Веса адаптивны к данным.

2. **Max Pooling** – классический метод подвыборки, который осуществляет уменьшение размерности карт признаков.
3. **WalsPooling** – это особый тип операции подвыборки, основанный на преобразовании Уолша. В отличие от стандартных методов (MaxPooling, AvgPooling), которые берут максимальное или среднее значение из региона, WalsPooling использует преобразование Уолша для агрегации признаков, что позволяет:
 - ◆ улучшить сжатие признаков с учетом их спектральных свойств;
 - ◆ сохранить больше информации, чем при стандартной подвыборке.

Алгоритм работы WalsPooling.

1. Применить прямое преобразование.

Данные получены после сверточного слоя. Размер передаваемой матрицы 256x256. Применяем прямое преобразование Уолша. Данная операция выполняется по следующей формуле:

$$\begin{aligned} \mathbf{W} &\in \{-1, +1\}^{256 \times 256} \\ \mathbf{Y} &= \mathbf{W} \cdot \mathbf{X} \cdot \mathbf{W}^T, \end{aligned} \quad (3)$$

где \mathbf{W} – матрица Уолша, \mathbf{X} – входной сигнал.

2. Фильтрация низкочастотных составляющих.

Применим фильтр низких частот, обнулив 75% значений спектральной составляющей (рис. 2,б,д).

Чтобы их обнулить, применим маску \mathbf{M} :

$$M_{ij} = \begin{cases} 0, & \text{если } i + j \leq K \\ 1, & \text{иначе} \end{cases}, \quad (4)$$

где K – порог обнуления.

Результирующая спектральная матрица

$$\mathbf{W}' = \mathbf{W} \cdot \mathbf{M}, \quad (5)$$

где \mathbf{M} – поэлементное умножение.

3. Применить обратное преобразование Уолша.

$$\mathbf{X}' = \frac{1}{256^2} \mathbf{W} \cdot \mathbf{W}' \cdot \mathbf{W}. \quad (6)$$

На рис. 2 приведена пара сравниваемых изображений, полученная с разных сенсоров: а – изображение с БПЛА, г – тайл подстилающей поверхности, б, д – фильтр низких частот, в, е – изображение после обратного преобразования.

4. Выполнить субдискретизацию.

После прямого преобразования и фильтрации производится обратное преобразование и субдискретизация (downsampling), выбирается только один элемент из каждого блока 2×2 . На рис. 3 приведена демонстрация субдискретизации в качестве операции подвыборки.

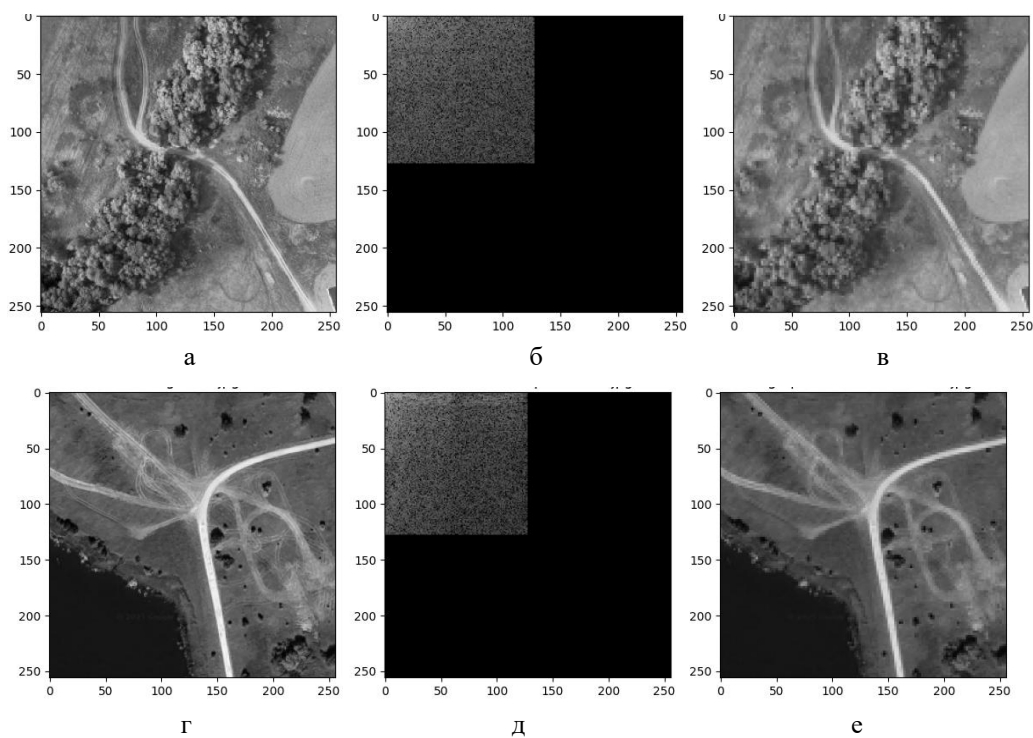


Рис. 2. Пара изображений после обратного преобразования Уолша

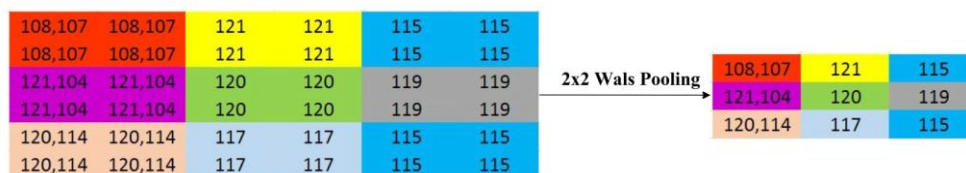


Рис. 3. Демонстрация операции подвыборки WalsPooling

5. Передать данные дальше по нейронной сети.

На рис. 4 приведена визуализация первого и второго слоя нейронной сети с различными операциями для уменьшения размерности. Для простоты восприятия в качестве входного изображения используем черно-белое изображение. На рис. 5 приведены гистограммы распределения значений, полученных после третьей свертки с применением преобразования Уолша для ядра свертки.

Преимуществом использования WalsPooling является быстрое вычисление, исходя из того, что матрица Уолша состоит только из -1 и +1.

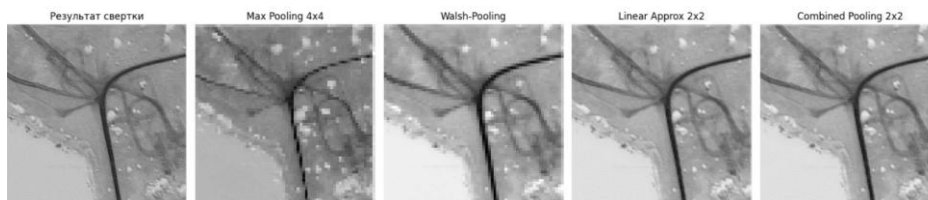


Рис. 4. Визуализация первого и второго слоя спектральной нейронной сети с различными операциями для уменьшения размерности

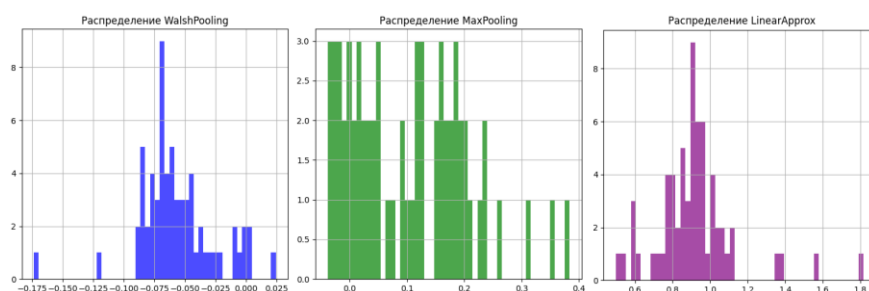


Рис. 5. Гистограммы распределения значений для различных операций уменьшения размерности

Исходя из рис. 5 видно, что гистограмма распределения значений для операции WalshPooling и для линейной аппроксимации к центру квадрата идентична. Ниже приведено сравнение характеристик выходов:

```
Сравнение характеристик выходов:
WalshPooling: min=0.7067, max=1.5253, mean=0.9676
MaxPooling: min=0.9746, max=3.2710, mean=1.7445
LinearApprox: min=0.4994, max=1.8163, mean=0.9187
```

```
Финальные предсказания:
WalshPooling: 0.650653
MaxPooling: 0.558279
LinearApprox: 0.279787
```

Однако, при схожей гистограмме распределений операций WalshPooling и LinearApprox финальные предсказания отличаются. Финальное предсказание – это выходы нейросети после прохождения всех слоев, включая пулинг и полносвязные слои, сжатые до одного числа в диапазоне $[0, 1]$ (благодаря сигмоиде в последнем слое). WalshPooling и LinearApprox имеют одинаковые min, max и mean (гистограммы близки), но финальные предсказания у них сильно различаются (0.6506, 0.2089).

- ◆ WalshPooling \rightarrow 0.6506.
- ◆ MaxPooling \rightarrow 0.5583.
- ◆ LinearApprox \rightarrow 0.279787.

Это означает, что

- ◆ WalshPooling даёт предсказания ближе к 1, что необходимо в задачах классификации.

- ◆ LinearApprox сильно занижает выход (возможно, из-за сглаживания).

- ◆ MaxPooling – промежуточный вариант, но ближе к WalshPooling.

Для того, чтобы улучшить результат, можно рассмотреть:

- ◆ комбинацию методов (WalshPooling + MaxPooling);
- ◆ изменить матрицу Уолша (например, использовать только низкочастотные компоненты).

Архитектура гибридной свёрточной нейронной сети. На вход нейронной сети подаются два изображения, полученные с разных сенсоров. На основе визуального сходства осуществляется сравнение пары изображений. Выходной слой свёрточной нейронной сети дает численную оценку близости пары изображений. В качестве операций для уменьшения размерности в слоях нейронной сети используется линейная интерполяция, подвыборка Max Pooling и разработанный WalsPooling. WalsPooling – это особый тип операции подвыборки, основанный на преобразовании Уолша. Рассмотрены гистограммы распределения значений после операций подвыборки.

На вход свёрточной нейронной сети подаются изображения, после предварительной обработки, уменьшенное до 256×256 . В качестве предварительной обработки применим спектральное преобразование Уолша к изображению, затем удалим низкочастотную составляющую и восстановим изображение [20–22].

На рис. 6 приведена модель предлагаемой сиамской нейронной сети.

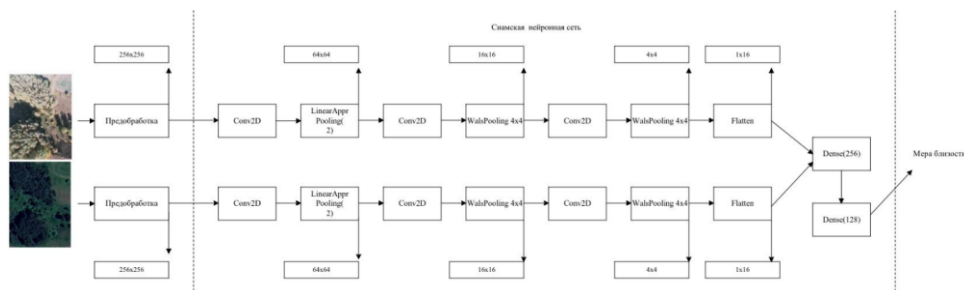


Рис. 6. Модель предлагаемой сиамской нейронной сети

Рассмотрим разработанную архитектуру свёрточной нейронной сети:

1. Исходное изображение 256×256 преобразуется в спектр Уолша, удаляются низкочастотные компоненты и восстанавливается изображение.
2. Применяем операция свертки к предобработанному изображению с ядром свёртки, вычисленным на основе обучения нейронной сети.
3. К полученной после операции свёртки 256×256 применяем линейную интерполяцию. Тем самым получаем 64×64 значения операции свёртки после усечения.
4. Применяем операцию свёртки 64×64 для полученной на шаге 3 сверки (после усечения).
5. К полученной после операции свёртки 64×64 применяем операцию усечения (WalsPooling 4×4), с помощью слоя нейронной сети, представленного в частотной области. Получим 16×16 значения операции свёртки после усечения.
6. Применяем к полученной после усечения WalsPooling операцию свёртки с ядром свёртки, вычисленным на основе обучения нейронной сети.
7. Применяем к полученной свертке 16×16 операцию усечения (WalsPooling 4×4). Получим 4×4 значений.
8. Полученные 4×4 значений преобразуем из многомерного пространства в одномерный вектор, чтобы его можно было подать на вход полносвязному слою (Dense) в нейронной сети. Получим 16 значений.
9. Следующие два полносвязных слоя осуществляют линейное преобразование + активацию. На первом полносвязном слое на вход поступает 16 нейронов с предыдущего слоя и 256 весов.

$$h_9 = \sigma(W_9 x_{flat} + b_9), \quad (7)$$

где x_{flat} – вектор, полученный на 8 шаге, $W_9 \in R^{256 \times 16}$ – матрица весов, $b_9 \in R^{256}$, $\sigma(z)$ – сигмоидная активация. Итого на выходе с данного слоя получим $4096 + 256 = 4352$ параметров.

10. Второй полносвязный слой Dense(128). На вход поступает 256 нейронов с предыдущего слоя. Итого на выходе $32768 + 128 = 32896$ параметров.

Dense(256) принимает 16 значений и выдаёт 256 значений $\rightarrow 4352$ параметра. Dense(128) принимает 256 значений и выдаёт 128 значений $\rightarrow 32,896$ параметров. Общее число параметров до сиамской части: 37,248.

11. После шагов 9 и 10 формируется две ветви (для изображений A и B):

$$x_{11}^a, x_{11}^b \in R^{128},$$

12. Общие слои для обеих ветвей. Повторим два полносвязных слоя Dense(256), Dense(128).

13. Сравнение ветвей осуществляется с использованием вычитания. Применяется квадратичная активация.

14. Финальный слой Dense(1). На вход подается вектор из 128 значения и 1 смещения. Получаем 129 параметров, которые описывают степень соответствия.

Полные формулы предложенной архитектуры

Для двух изображений A и B:

1. Спектральные свертки и подвыборка:

$$\begin{aligned} x_{flat}^a &= Flatten(WalshPooling(Conv3(Walsh(A)))) \\ x_{flat}^b &= Flatten(WalshPooling(Conv3(Walsh(B)))) \end{aligned} \quad (8)$$

2. Общие полносвязные слои:

$$\begin{aligned} x_{11}^a &= \sigma(W_{10}(\sigma(W_9 x_{flat}^a + b_9) + b_{10})) \\ x_{11}^b &= \sigma(W_{10}(\sigma(W_9 x_{flat}^b + b_9) + b_{10})). \end{aligned} \quad (9)$$

3. Сиамская сеть:

$$y = w_{out} [\sigma(W_{14}\sigma(W_{13}x_{11}^a + b_{13}) + b_{14}) - \sigma(W_{14}\sigma(W_{13}x_{11}^b + b_{13}) + b_{14})]^2 + b_{out}. \quad (10)$$

На рис. 5,6 приведена визуализация работы предложенной архитектуры гибридной свёрточной нейронной сети. Различия зависят от ядра свёртки, которое формируется в процессе обучения нейронной сети.

Визуализация слоев нейронной сети. На рис. 7 и 8 приведена визуализация слоев для пары изображений, полученных с разных сенсоров. На рис. 7, 8:

- а) – результат после первой операции свёртки;
- б) – результат после операции уменьшения размерности (линейная аппроксимация к центру квадрата);
- в) – результат после второй операции свёртки;
- г) – результат после подвыборки разработанным WalshPooling;
- д) – результат после операции свёртки;
- е) – результат после подвыборки разработанным WalshPooling.

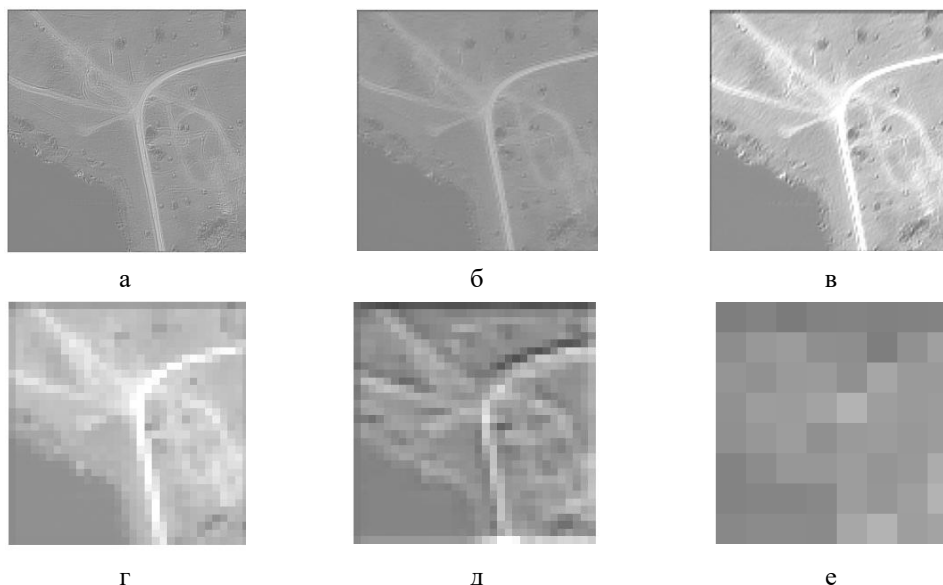


Рис. 7. Визуализация работы 6 слоев гибридной свёрточной нейронной сети для снимка с БПЛА

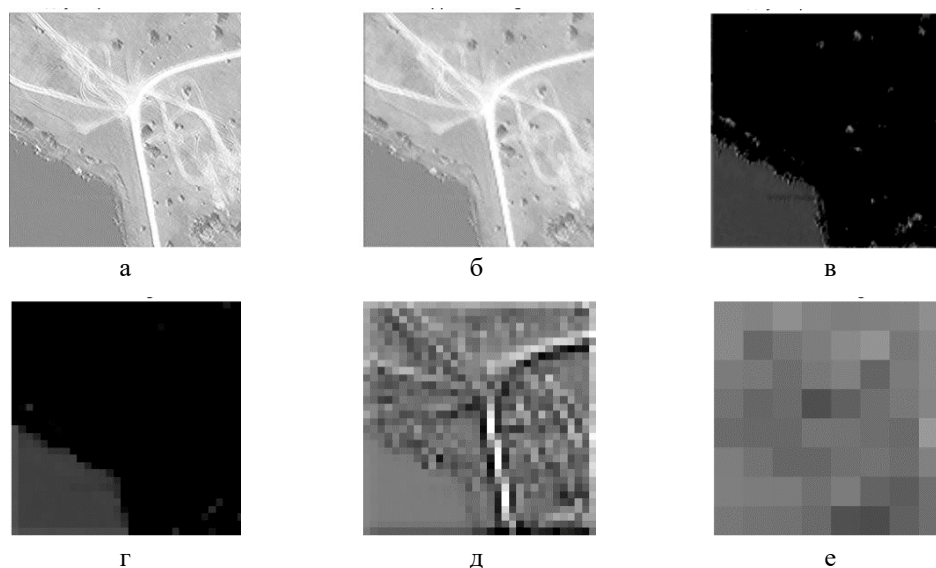


Рис. 8. Визуализация работы 6 слоев гибридной свёрточной нейронной сети для тайла подстилающей поверхности

В данной архитектуре отсутствуют активационные функции, что позволяет снизить размерность слоев нейронной сети, а также ускоряет сборку и обучение нейронной сети. Однако, можно добавить в разработанную архитектуру активационные функции и сравнить сходимости сети с ними и без них.

Подход к формированию набора данных для обучения и результаты экспериментов. Проведем эксперимент на разработанной архитектуре. В данной статье предложен подход к формированию данных для обучения нейронной сети. В качестве искажений применяются следующие виды трансформаций (табл. 2).

Таблица 2

Виды трансформации

Категория трансформации	Тип трансформации	Базовое значение и разброс
Яркостные искажения	Изменение яркости	$1.0 \pm 0.3\%$
	Изменение контрастности	$1.0 \pm 0.3\%$
	Добавление Гауссова шума	0.0 ± 0.2
Геометрические искажения	Поворот изображения	$0.0 \pm 5.0^\circ$
	Масштабирование	1.5 ± 0.25

Сравнение результатов работы сходимости нейронной сети производилось на:

- ◆ парах изображений без добавления шума и аугментации;
- ◆ парах изображений с добавлением шума и аугментации;
- ◆ парах изображений после применения спектрального преобразования Уолша и восстановленного изображения;
- ◆ парах изображений после применения спектрального преобразования Уолша и восстановленного изображения без учета постоянной составляющей;
- ◆ парах изображений после применения спектрального преобразования Уолша, усеченных на 25%, 50%, 75% значений;
- ◆ парах изображений восстановленного изображения, усеченных на 25%, 50%, 75% значений;
- ◆ парах изображений после спектральных преобразований с предварительным добавлением шума, аугментации и поворота;

В табл. 3 приведены результат работы сиамской нейронной сети.

Обозначения:

1. Без искажений.
2. Яркостные искажения.
3. Яркостные и геометрические искажения.
 - a) Без удаления информационной составляющей.
 - b) С удалением 25%.
 - c) С удалением 50%.
 - d) С удалением 75%.

x – Validation Accuracy (пары спутниковых снимков в разных локациях с искажениями).

y – Validation Loss (сред. квадр. ошибка).

z – Test Accuracy (пары спутник/беспилотник в идентичных локациях без искажений).

Таблица 3

Результат работы сиамской нейронной сети с разным набором входных значений

Тип входных данных	Спектр Уолша			Спектр Уолша без постоянной составляющей			Восстановленное изображение			Восстановленное изображение без постоянной составляющей		
	x	y	z	x	y	z	x	y	z	x	y	z
Метрика												
A-1	0.4975	2.4463	1	0.5013	2.2463	1	0.7494	1.1650	7	0.7400	1.0977	5
A-2	0.5088	2.6302	1	0.4937	2.4682	1	0.7800	1.0922	6	0.7325	1.0530	5
A-3	0.5075	2.5378	1	0.4900	3.0640	1	0.6875	1.4485	4	0.6784	1.5660	5
B-1	0.5300	2.3623	2	0.4586	2.3105	1	0.8175	0.7654	5	0.7600	1.2870	5
B-2	0.5000	2.6127	1	0.4727	2.4101	1	0.7675	0.5376	7	0.7800	0.8082	6
B-3	0.5125	2.5174	1	0.4782	2.4473	0	0.7750	0.7742	6	0.7118	1.1856	7
C-1							0.7085	0.8285	6	0.7775	1.2558	4
C-2							0.7425	0.9872	7	0.7450	1.1599	6
C-3							0.7250	0.8841	5	0.7425	1.1982	7
D-1							0.7500	1.1539	8	0.7025	1.1021	4
D-2							0.7965	0.9484	6	0.7100	1.2198	6
D-3							0.7450	0.8754	4	0.7425	1.2830	6

Низкий показатель **x** – Validation Accuracy (пары спутниковых снимков в разных локациях с искажениями). **y** – Validation Loss (сред. квадр. ошибка) на наборах данных, где в качестве входного изображения выступает спектральное представление связано с тем, что сверточная нейронная сеть в операциях свёртки оперирует не частотным представлением, поэтому необходимо переходить от пространственного представления к частотному. Наилучший результат СНС показывает в том случае, если в качестве входного слоя применяется восстановленное изображение, причем можно убирать до 75% значений, в том числе и постоянную составляющую. Лучшая конфигурация: Стандартные свёртки с восстановлением дают максимальную Validation Accuracy (0.7965) и приемлемый Loss (0.9484).

Метрики информационного поиска. Для того, чтоб оценить точность поиска рассмотрим следующие метрики:

♦ точность поиска (precision) вычисляется как отношение выданных релевантных изображений к общему количеству выданных изображений;

$$p = \frac{RF}{RF+IF}, \tag{11}$$

где RF – количество выданных релевантных изображений, IF – количество выданных нерелевантных изображений.

♦ полнота поиска (recall) вычисляется как отношение выданных релевантных изображений к общему числу релевантных изображений;

$$r = \frac{RF}{RF+RN}, \quad (12)$$

где RN – количество релевантных не выданных значений.

F1-мера (сбалансированная F-мера, мера Ван Ризбергена, F1 – measure) применяется как единая мера, объединяющая полноту и точность, которая вычисляется следующим образом:

$$F_1 = \frac{2rp}{r+p}. \quad (13)$$

F1-мера (значение F1-меры не может быть ниже полноты и точности и превышать их среднеарифметическое):

$$\min(r, p) \leq F_1 \leq \frac{r+p}{2}; \quad (14)$$

Для предложенной архитектуры получим следующие результаты, приведенные в табл. 4, 5.

Таблица 4

Метрики информационного поиска для предложенной архитектуры

Метрика	Значение	Интерпретация
Precision	0.728	72.8% пар, которые модель назвала "похожими", действительно похожи
Recall	0.800	Модель находит лишь 80.0% всех реально похожих пар
F1	0.872	Значение 0.872 близко к 1, что указывает на высокое качество модели в балансе между: precision (точностью) и recall (полнотой).
AUC-ROC	0.784	Значение 0.784 означает, что в 78.4% случаев модель верно ранжирует случайно выбранный положительный пример выше, чем отрицательный.

Таблица 5

Средние расстояния между парами

Класс	Значение	Интерпретация
Same (похожие)	0,317	Среднее расстояние для пар одной локации
Diff (разные)	1.869	Среднее расстояние для пар разных локаций

Заключение. В работе предложена гибридная архитектура СНС, сочетающая спектральные и пространственные слои для оценки географической близости изображений, полученных с разных сенсоров: спутниковое изображение и изображение, полученное с БПЛА. Введены новые методы подвыборки и свёртки, демонстрирующие устойчивость к изменениям масштаба и угла съёмки. Использована сиамская архитектура для обработки пар изображений, что критично для задач сопоставления. В дальнейшем предлагается расширить набор данных для обучения нейронной сети с разными условиями съёмки (погода, время суток) для покрытия большего числа сценариев, ввести дополнительные метрики IoU (Intersection over Union) для оценки геометрического перекрытия предсказанных и истинных областей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Gaihua Wang, Guoliang Yuan, Meng Lv, WenZhou Liu.* Application of Interpolation Pooling in Convolutional Neural Networks // *Helix.* – 2018. – Vol. 8 (4). – P. 3465- 3469. – DOI 10.29042/2018-3465-3469
2. *Zhuo Chen, Jiyuan Zhang, Ruizhou Ding.* ViP: Virtual Pooling for Accelerating CNN-based Image Classification and Object Detection // *WACV.* – 2020 – P. 1180-1189. – https://openaccess.thecvf.com/content_WACV_2020/papers/Chen_ViP_Virtual_Pooling_for_Accelerating_CNN-based_Image_Classification_and_Object_WACV_2020_paper.pdf.

3. *Travis Williams, Robert Li.* Wavelet Pooling for Convolutional Neural Networks // Published as a conference paper at ICLR. – 2018. – P. 1-12. – <https://openreview.net/forum?id=rkhlb8ICZ>.
4. *Juan Manuel Fortuna-Cervantes, Marco Tulio Ramirez-Torres, Marcela Mejía-Carlos, José Salomé Murguía, José Martínez-Carranza, Carlos Soubervielle-Montalvo, César Arturo Guerra-García.* Texture and Materials Image Classification Based on Wavelet Pooling Layer in CNN // the Special Issue Recent Advances in Deep Learning for Image Analysis. – <https://doi.org/10.3390/app12073592>.
5. Image-Net. – <https://www.image-net.org/>.
6. *Pengju Liu, Hongzhi Zhang, Wei Lian, Wangmeng Zuo.* Multi-Level Wavelet Convolutional Neural Networks // Digital Object Identifier. – 10.1109/ACCESS.2019.2921451. – <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8732332>.
7. *Вершков Н.А., Бабенко М.Г., Черных А.Н., Кучуков В.А., Кучеров Н.Н., Кучукова Н.Н.* Оптимизация искусственных нейронных сетей с помощью вейвлет-преобразований // Программирование. – 2022. – № 6. – С. 22-30. – DOI: 10.31857/S0132347422060073.
8. *Spitsyn V.G., Bolotova Yu.A., Phan N.H., Bui T.T.T.* Using a Haar wavelet transform, principal component analysis and neural networks for OCR in the presence of impulse noise. – <https://computeroptics.ru/eng/KO/Annot/KO40-2/400216.html>.
9. *Митина О.А., Базанова Д.В.* Разработка модели нейронной сети на основе ряда Фурье // НАУ. – 2023. – № 90-2. – URL: <https://cyberleninka.ru/article/n/razrabotka-modeli-neyronnoy-seti-na-osnove-ryada-furje> (дата обращения: 12.08.2025).
10. *Фортуна-Сервантес Х.М., Рамирес-Торрес М.Т., Мартинес-Карранса Х., Мургуиа-Ибарра Х.С., Мехиа-Карлос М.* Обнаружение объектов в аэронавигации с использованием вейвлет-преобразования и сверточных нейронных сетей: первый подход // Тр. ИСП РАН. – 2021. – Т. 33. – Вып. 2. – DOI: 10.15514/ISPRAS-2020-33(2)-9.
11. *Беляев П.Ю., Зикратов И.А.* Исследование автономной навигации беспилотных летательных аппаратов на основе корреляционных методов сравнения изображений // Тр. учебных заведений связи. – 2024. – Т. 10, № 5. – С. 109-118.
12. *Давыдова А.А., Чечиков Ю.Б.* Проблема реализации современной GPS-навигации для БПЛА // Перспективные научные исследования: опыт, проблемы и перспективы развития: Сб. научных статей по материалам X Международной научно-практической конференции. – Уфа, 2023. – С. 100-106.
13. *Кориков А.М., Тран В.Т.* Применение нейронных сетей и корреляционно-экстремального координатора для навигации и управления БПЛА // Природные и интеллектуальные ресурсы Сибири (СИБРЕСУРС-28-2022): Доклады (материалы конференции) 28-й международной научно-практической конференции. – Томск, 2022. – С. 29-33.
14. *Лесив Е.А., Вытовтов К.А., Барабанова Е.А.* Нейросетевой алгоритм анализа видеoinформации для резервной системы локальной навигации привязных БПЛА с обучением на синтетических данных // Информационные технологии и технические средства управления (ИСТТ- 2023). Матер. VII Международной научной конференции. – М., 2023. – С. 159-162.
15. *Ustinova V.E., Lutsenko A.S., Shpak A.V., Mironenkov G.V., Ivlev V.A.* A method for finding the correspondence between a railway station model and its visual representation based on graphs // Computing, Telecommunications and Control. – 2024. – Vol. 17, No. 4. – P. 64-77.
16. *Ustinova V.E., Lutsenko A.S., Shpak A.V., Mironenkov G.V., Ivlev V.A.* A method for finding the correspondence between a railway station model and its visual representation based on graphs // Computing, Telecommunications and Control. – 2024. – Vol. 17, No. 4. – P. 64-77.
17. *Кочетунов В.В., Федяев О.И.* Сиамские сверточные нейронные сети для распознавания лиц // Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПШИВС-2024): Сб. материалов и докладов V Международной научно-практической конференции. – Донецк, 2024. – С. 178-184.
18. *Ладанова Е.О.* Разработка методики и новой структуры сиамской нейронной сети для анализа геопространственных данных // Научно-технический вестник Поволжья. – 2024. – № 12. – С. 144-147.
19. *Sumina E.A.* Siamese networks as a possible improvement of standard algorithms // Artificial Intelligence: Challenges, Essence, Communication, Creativity and other aspects: Сб. материалов студенческих конференций. – М., 2024. – С. 104-108.
20. *Куртяник Д. В., Сергеев А.М., Самиков А.В., Семенов А.А.* О визуальной оценке результатов поиска матриц Адамара // Наука. Техника. Технологии (политехнический вестник). – 2018. – № 4. – С. 19-26.
21. *Кузков С.И., Цветков А.В.* Основы теории сигналов: учеб. пособие. – Екатеринбург: Изд-во Урал. ун-та, 2023. – 114 с. – ISBN 978-5-7996-3636-4.
22. *Костров Б.В., Гринченко Н.Н., Баранова С.Н., Трушина Е.А., Вьюгина А.А.* Ортогональное кодирование бинарных изображений // Вестник Ярославского высшего военного училища противовоздушной обороны. – 2023. – № 2 (21). – С. 82-87.

REFERENCES

1. *Gaihua Wang, Guoliang Yuan, Meng Lv, WenZhou Liu.* Application of Interpolation Pooling in Convolutional Neural Networks, *Helix*, 2018, Vol. 8 (4), pp. 3465- 3469. DOI 10.29042/2018-3465-3469
2. *Zhuo Chen, Jiyuan Zhang, Ruizhou Ding.* ViP. Virtual Pooling for Accelerating CNN-based Image Classification and Object Detection, *WACV*, 2020, pp. 1180-1189. Available at: https://openaccess.thecvf.com/content_WACV_2020/papers/Chen_ViP_Virtual_Pooling_for_Accelerating_CNN-based_Image_Classification_and_Object_WACV_2020_paper.pdf.
3. *Travis Williams, Robert Li.* Wavelet Pooling for Convolutional Neural Networks, *Published as a conference paper at ICLR*, 2018, pp. 1-12. Available at: <https://openreview.net/forum?id=rkhlb8ICZ>.
4. *Juan Manuel Fortuna-Cervantes, Marco Tulio Ramirez-Torres, Marcela Mejia-Carlos, José Salomé Murguía, José Martínez-Carranza, Carlos Soubervielle-Montalvo, César Arturo Guerra-García.* Texture and Materials Image Classification Based on Wavelet Pooling Layer in CNN, *the Special Issue Recent Advances in Deep Learning for Image Analysis*. Available at: <https://doi.org/10.3390/app12073592>.
5. Image-Net. Available at: <https://www.image-net.org/>.
6. *Pengju Liu, Hongzhi Zhang, Wei Lian, Wangmeng Zuo.* Multi-Level Wavelet Convolutional Neural Networks, *Digital Object Identifier*. 10.1109/ACCESS.2019.2921451. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8732332>.
7. *Vershkov N.A., Babenko M.G., Chernykh A.N., Kuchukov V.A., Kucherov N.N., Kuchukova N.N.* Optimizatsiya iskusstvennykh neyronnykh setey s pomoshch'yu veyvlet-preobrazovaniy [Optimization of artificial neural networks using wavelet transforms], *Programirovanie [Programming]*, 2022, No. 6, pp. 22-30. DOI: 10.31857/S0132347422060073.
8. *Spitsyn V.G., Bolotova Yu.A., Phan N.H., Bui T.T.T.* Using a Haar wavelet transform, principal component analysis and neural networks for OCR in the presence of impulse noise. Available at: <https://computeroptics.ru/eng/KO/Annot/KO40-2/400216.html>.
9. *Mitina O.A., Bazanova D.V.* Razrabotka modeli neyronnoy seti na osnove ryada Fur'e [Development of a neural network model based on the Fourier series], *NAU [National Association of Scientists]*, 2023, No. 90-2. Available at: <https://cyberleninka.ru/article/n/razrabotka-modeli-neyronnoy-seti-na-osnove-ryada-furie> (accessed 12 August 2025).
10. *Fortuna-Servantes Kh.M., Ramires-Torres M.T., Martines-Karransa Kh., Murguía-Ibarra Kh.S., Mekhia-Karlos M.* Obnaruzhenie ob"ektov v aeronavigatsii s ispol'zovaniem veyvlet-preobrazovaniya i svetochnykh neyronnykh setey: pervyy podkhod [Object Detection in Air Navigation Using Wavelet Transform and Convolutional Neural Networks: A First Approach], *Tr. ISP RAN [Proceedings of ISP RAS]*, 2021, Vol. 33, Issue 2. DOI: 10.15514/ISPRAS-2020-33(2)-9.
11. *Belyaev P.Yu., Zikratov I.A.* Issledovanie avtonomnoy navigatsii bespilotnykh letatel'nykh apparatov na osnove korrelyatsionnykh metodov sravneniya izobrazheniy [Study of Autonomous Navigation of Unmanned Aerial Vehicles Based on Correlation Methods of Image Comparison], *Tr. uchebnykh zavedeniy svyazi [Proceedings of Communications Educational Institutions]*, 2024, Vol. 10, No. 5, pp. 109-118.
12. *Davydova A.A., Chechikov Yu.B.* Problema realizatsii sovremennoy GPS-navigatsii dlya BPLA [The problem of implementing modern GPS navigation for UAVs], *Perspektivnye nauchnye issledovaniya: opyt, problemy i perspektivy razvitiya: Sb. nauchnykh statey po materialam X Mezhdunarodnoy nauchno-prakticheskoy konferentsii [Advanced scientific research: experience, problems and development prospects: Collection of scientific articles based on the materials of the X International scientific and practical conference]*. Ufa, 2023, pp. 100-106.
13. *Korikov A.M., Tran V.T.* Primenenie neyronnykh setey i korrelyatsionno-ekstremalnogo koordinatora dlya navigatsii i upravleniya BPLA [Application of neural networks and correlation-extreme coordinator for navigation and control of UAVs], *Prirodnye i intellektual'nye resursy Sibiri (SIBRESURS-28-2022): Doklady (materialy konferentsii) 28-y mezhdunarodnoy nauchno-prakticheskoy konferentsii [Natural and intellectual resources of Siberia (SIBRESURS-28-2022): Reports (conference materials) of the 28th international scientific and practical conference]*. Tomsk, 2022, pp. 29-33.
14. *Lesiv E.A., Vytovtov K.A., Barabanova E.A.* Neyrosetevoy algoritm analiza videoinformatsii dlya rezervnoy sistemy lokal'noy navigatsii privyaznykh BPLA s obucheniem na sinteticheskikh dannykh [Neural network algorithm for analyzing video information for a backup local navigation system for tethered UAVs with training on synthetic data], *Informatsionnye tekhnologii i tekhnicheskie sredstva upravleniya (ICCT- 2023). Mater. VII Mezhdunarodnoy nauchnoy konferentsii [Information technology and technical means of control (ICCT-2023). Proceedings of the VII International Scientific Conference]*. Moscow, 2023, pp. 159-162.
15. *Ustinova V.E., Lutsenko A.S., Shpak A.V., Mironenkov G.V., Ivlev V.A.* A method for finding the correspondence between a railway station model and its visual representation based on graphs, *Computing, Telecommunications and Control*, 2024, Vol. 17, No. 4, pp. 64-77.

16. Ustinova V.E., Lutsenko A.S., Shpak A.V., Mironenkov G.V., Ivlev V.A. A method for finding the correspondence between a railway station model and its visual representation based on graphs, *Computing, Telecommunications and Control*, 2024, Vol. 17, No. 4, pp. 64-77.
17. Kocheturov V.V., Fedyaev O.I. Siamskie svertochnye neyronnye seti dlya raspoznavaniya lits [Siamese convolutional neural networks for face recognition], *Programmaya inzheneriya: metody i tekhnologii razrabotki informatsionno-vychislitel'nykh sistem (PIIVS-2024): Sb. materialov i dokladov V Mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Software Engineering: Methods and Technologies for Developing Information and Computing Systems (PIIVS-2024): Collection of materials and reports of the V International scientific and practical conference]. Donetsk, 2024, pp. 178-184.
18. Ladanova E.O. Razrabotka metodiki i novoy struktury siamskoy neyronnoy seti dlya analiza geoprostranstvennykh dannyykh [Development of a methodology and a new structure of a Siamese neural network for analyzing geospatial data], *Nauchno-tekhnicheskii vestnik Povolzh'ya* [Scientific and Technical Bulletin of the Volga Region], 2024, No. 12, pp. 144-147.
19. Sumina E.A. Siamese networks as a possible improvement of standard algorithms, *Artificial Intelligence: Challenges, Essence, Communication, Creativity and other aspects: Collection of materials from student conferences*. Moscow, 2024, pp. 104-108.
20. Kurtyanik D. V., Sergeev A.M., Samikov A.V., Semenov A.A. O vizual'noy otsenke rezul'tatov poiska matrits Adamara [On the visual assessment of the results of the search for Hadamard matrices], *Nauka. Tekhnika. Tekhnologii (politekhniceskii vestnik)* [Science. Engineering. Technologies (Polytechnic Bulletin)], 2018, No. 4, pp. 19-26.
21. Kumkov S.I., TSvetkov A.V. Osnovy teorii signalov: ucheb. posobie [Fundamentals of signal theory: a tutorial]. Ekaterinburg: Izd-vo Ural. un-ta, 2023, 114 p. ISBN 978-5-7996-3636-4.
22. Kostrov B.V., Grinchenko N.N., Baranova S.N., Trushina E.A., V'yugina A.A. Ortogonal'noe kodirovaniye binarnyykh izobrazheniy [Orthogonal coding of binary images], *Vestnik Yaroslavskogo vysshogo voennogo uchilishcha protivovozdushnoy oborony* [Bulletin of the Yaroslavl Higher Military School of Air Defense], 2023, No. 2 (21), pp. 82-87.

Костров Борис Васильевич – Рязанский государственный радиотехнический университет; e-mail: kostrov.b.v@evm.rsreu.ru; г. Рязань, Россия; тел.: +79036930570; профессор; зав. кафедрой электронных вычислительных машин.

Бабаев Сергей Игоревич – Рязанский государственный радиотехнический университет; e-mail: babaev.s.i@gmail.com; г. Рязань, Россия; тел.: +79209676677; доцент кафедры электронных вычислительных машин.

Ефимов Алексей Игоревич – Рязанский государственный радиотехнический университет; e-mail: lexie62rus@mail.ru; г. Рязань, Россия; тел.: +79105665083; доцент кафедры электронных вычислительных машин.

Тарасова Валентина Юрьевна – Рязанский государственный радиотехнический университет; e-mail: Valentina2008.91@mail.ru; тел.: +79038373769; старший преподаватель кафедры электронных вычислительных машин.

Kostrov Boris Vasilyevich – Ryazan State Radio Engineering University; e-mail: kostrov.b.v@evm.rsreu.ru; Ryazan, Russia; phone: +79036930570; professor, head of Department of Electronic Computers.

Babaev Sergey Igorevich – Ryazan State Radio Engineering University; e-mail: babaev.s.i@gmail.com; Ryazan, Russia; phone: +79209676677; associate Professor, Department of Electronic Computers.

Efimov Alexey Igorevich – Ryazan State Radio Engineering University; e-mail: lexie62rus@mail.ru; Ryazan, Russia; phone: +79105665083; associate professor, Department of Electronic Computers.

Tarasova Valentina Yurievna – Ryazan State Radio Engineering University; e-mail: Valentina2008.91@mail.ru; Ryazan, Russia; phone: +79038373769; senior Lecturer, Department of Electronic Computers Machines.