

DOI 10.58351/2949-2041.2026.36.7.012

Сироткин Максим Сергеевич, Студент
Сахалинский государственный университет
Sirotkin Maxim Sergeevich, Sakhalin State University

Научный руководитель:
Осипов Геннадий Сергеевич
д.т.н., профессор кафедры информатики
Сахалинский государственный университет
Osipov Gennady Sergeevich, Sakhalin State University

**ГИБРИДНАЯ СИСТЕМА ОЦЕНКИ КАЧЕСТВА РУКОПИСНОГО ВВОДА
НА ОСНОВЕ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ОБУЧЕНИЯ
НАПИСАНИЮ БУКВ АНГЛИЙСКОГО АЛФАВИТА
HYBRID HANDWRITING QUALITY ASSESSMENT SYSTEM BASED
ON CONVOLUTIONAL NEURAL NETWORK FOR LEARNING
TO WRITE LETTERS OF ENGLISH ALPHABET**

Аннотация. В условиях цифровизации образования всё более актуальной становится задача создания обучающих систем, способных не только распознавать рукописный ввод, но и объективно оценивать его качество. Большинство существующих цифровых решений либо ограничиваются механической обводкой по шаблону, либо используют стандартные технологии оптического распознавания символов, которые недостаточно адаптированы к высокой вариативности детского почерка.

В работе представлена гибридная система оценки качества рукописного ввода, объединяющая свёрточную нейронную сеть, обученную на расширенном наборе данных EMNIST ByClass с добавлением синтетических образцов шрифта Caveat, и модуль постобработки, реализующий геометрический анализ изображения и эвристические правила. Такой подход позволяет одновременно определять введённый символ и количественно оценивать правильность его написания.

Научная новизна состоит в практической реализации гибридного метода, сочетающего возможности глубокого обучения с символическим анализом геометрических характеристик почерка. Разработанная модель демонстрирует точность до 99,177 % на данных, приближенных к реальному детскому почерку, при среднем времени инференса менее 500 миллисекунд. Web-приложение внедрено в образовательный процесс Детского технопарка «Кванториум» и обеспечивает сохранение результатов, выявление типичных ошибок и мониторинг индивидуального прогресса учащихся.

Abstract. In the context of digitalization of education, the task of creating training systems that can not only recognize handwritten input, but also objectively assess its quality is becoming more and more urgent. Most existing digital solutions are either limited to mechanical strokes on a template, or use standard optical character recognition technologies that are not sufficiently adapted to the high variability of children's handwriting.

The paper presents a hybrid handwriting quality assessment system that combines a convolutional neural network trained on the EMNIST ByClass extended dataset with the addition of synthetic Caveat font samples, and a post-processing module that implements geometric image analysis and heuristic rules. This approach allows you to simultaneously determine the entered symbol and quantify the correctness of its spelling.

The scientific novelty consists in the practical implementation of a hybrid method that combines the possibilities of deep learning with a symbolic analysis of the geometric characteristics of handwriting. The developed model demonstrates accuracy up to 99.177% on data close to real children's handwriting, with an average inference time of less than 500 milliseconds. The web application is implemented in the educational process of the Quantorium Children's Technopark and ensures the preservation of results, identification of typical errors and monitoring of individual progress of students.



Ключевые слова: Свёрточные нейронные сети; распознавание рукописного текста; гибридная система оценки качества; геометрический анализ; обучение письму; EMNIST ByClass; web-приложение.

Keywords: Convolutional neural networks; handwriting recognition; hybrid quality assessment system; Geometric analysis writing training; EMNIST ByClass; web application.

Введение

В последние годы цифровизация образовательного процесса привела к появлению большого количества обучающих систем, которые позволяют повысить эффективность формирования навыков чтения и письма у детей. Традиционные методы обучения письму, основанные на использовании печатных прописей и постоянном контроле со стороны педагога, имеют существенные ограничения: они не обеспечивают мгновенной обратной связи и плохо учитывают индивидуальные особенности детского почерка. Большинство существующих цифровых приложений либо ограничиваются механической обводкой букв по заданному шаблону, либо используют стандартные технологии оптического распознавания символов, которые недостаточно адаптированы к высокой вариативности размера, пропорций и положения символов относительно базовых линий, характерной для детей младшего школьного возраста.

В связи с этим целью работы стала разработка web-приложения, реализующего гибридную систему оценки качества рукописного ввода на основе свёрточной нейронной сети и геометрического анализа. Система сочетает возможности глубокого обучения с модулем постобработки, опирающимся на геометрические характеристики изображения и эвристические правила. Такой подход позволяет не только распознавать введённый символ, но и объективно оценивать правильность и разборчивость его написания.

Научная новизна исследования заключается в практической реализации гибридного метода оценки качества рукописного ввода. Практическая значимость работы состоит в создании инструмента, внедрённого в образовательный процесс Детского технопарка «Кванториум».

1. Теоретические основы машинного обучения и свёрточных нейронных сетей

1.1. Машинное обучение и нейронные сети

Машинное обучение представляет собой подход к анализу данных, при котором модель выявляет закономерности в данных и использует их для решения задач без явно заданных правил [7]. В отличие от традиционных алгоритмов, модель самостоятельно формирует зависимости между входными данными и целевыми значениями в процессе обучения на примерах.

Методы машинного обучения принято разделять на обучение с учителем, без учителя и с подкреплением. В задачах распознавания образов, включая распознавание рукописного текста, преимущественно применяется обучение с учителем [5]. В этом случае модель обучается на размеченных данных, где каждому изображению соответствует метка класса, что позволяет ей обобщать знания и корректно классифицировать новые примеры даже при значительной вариативности почерка.

Нейронные сети состоят из последовательности слоёв, в которых нейроны вычисляют взвешенную сумму входных сигналов с последующим применением нелинейной функции активации [1].

Полносвязные нейронные сети соединяют каждый нейрон со всеми нейронами следующего слоя. При работе с изображениями они имеют существенный недостаток – большое количество параметров, что приводит к высокой вычислительной сложности и риску переобучения.

Рекуррентные нейронные сети содержат обратные связи, позволяющие учитывать контекст предыдущих состояний. Они эффективны для последовательных данных, однако менее приспособлены для задач, где важна пространственная структура изображения.



Свёрточные нейронные сети специально разработаны для обработки изображений. Благодаря использованию операции свёртки с общими весами они значительно сокращают количество параметров, обеспечивают устойчивость к небольшим сдвигам и искажениям и позволяют автоматически извлекать иерархические признаки [4]. Именно эти свойства сделали свёрточные нейронные сети оптимальным выбором для распознавания рукописных символов.

1.2. Свёрточные нейронные сети: архитектура, компоненты и принципы работы

Свёрточные нейронные сети представляют собой специализированный класс архитектур, разработанный для эффективной обработки изображений. Их ключевое преимущество заключается в способности автоматически извлекать иерархические признаки с использованием относительно небольшого числа параметров, а также в устойчивости к небольшим изменениям положения и ориентации объектов. В контексте задачи распознавания рукописного текста эти свойства особенно ценны, поскольку детский почерк характеризуется высокой вариативностью формы, размера и положения символов.

1.2.1. Общая идея и архитектура свёрточной нейронной сети

Типичная архитектура свёрточной нейронной сети для классификации рукописных символов последовательно сочетает операции свёртки, подвыборки и полносвязных преобразований [6]. Как показано на рисунке 1, входное изображение размером 28×28 пикселей с одним каналом интенсивности проходит через два свёрточных слоя, после каждого из которых применяется операция Max-Pooling с окном 2×2 . Результаты свёрток разворачиваются в одномерный вектор и подаются на полносвязные слои с функцией активации ReLU и Dropout, завершающиеся выходным слоем с функцией Softmax.

Данная схема носит иллюстративный характер и адаптирована под 62 класса (заглавные и строчные буквы английского алфавита вместе с цифрами), которые используются в реальной модели web-приложения.

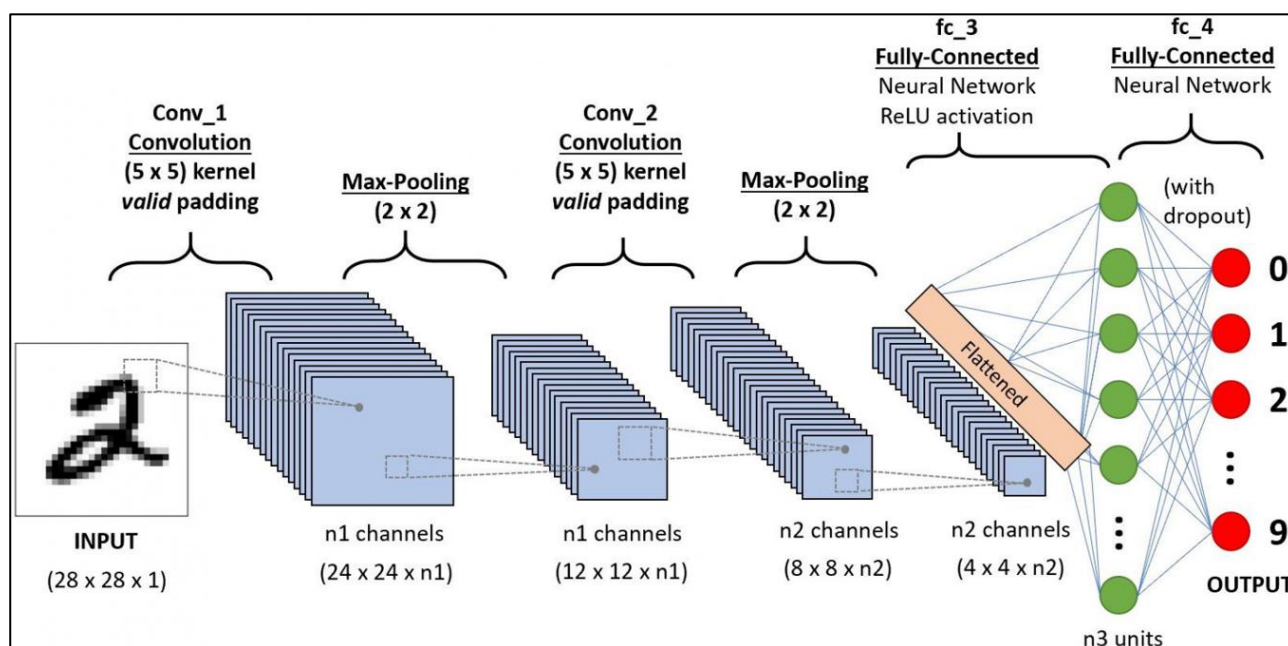


Рисунок 1 – Общая архитектура свёрточной нейронной сети для классификации рукописных цифр

В разрабатываемом web-приложении применяется архитектура, близкая к представленной схеме, но адаптированная под задачу классификации 62 символов. Модель состоит из двух свёрточных слоёв с последующими операциями подвыборки, полносвязного слоя на 128 нейронов и выходного слоя на 62 класса. Такой выбор позволяет сохранять баланс между выразительной способностью сети и вычислительной эффективностью.

1.2.2. Входное изображение и его представление в виде тензора

Любое изображение, подаваемое на вход свёрточной нейронной сети, может быть представлено в виде многомерного массива чисел, называемого тензором. На рисунке 2 показан пример RGB-изображения размером 4×4 пикселя, которое кодируется как тензор размерности 4×4×3.

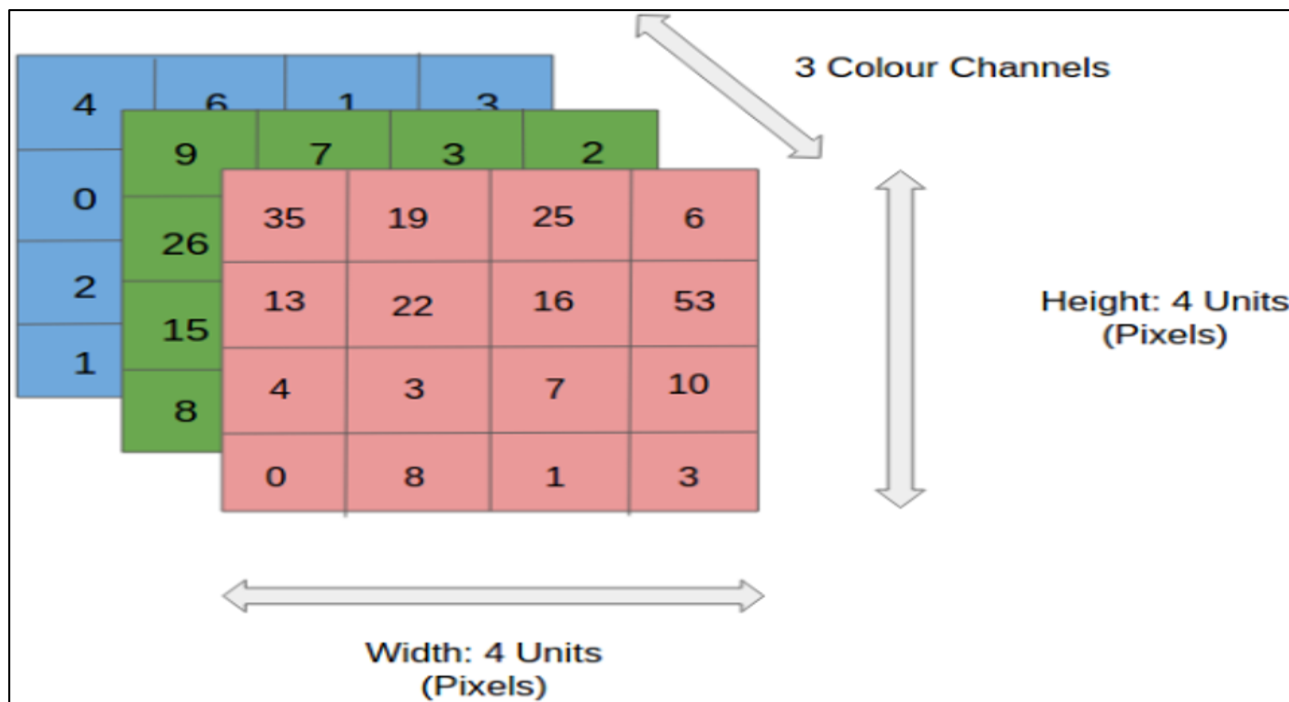


Рисунок 2 – Представление RGB-изображения как тензора размером 4×4×3

В случае с наборами данных EMNIST ByClass и синтетическими образцами, сгенерированными шрифтом Caveat, изображения преобразуются в одноканальный формат (grayscale). Это объясняется тем, что для задач распознавания рукописного текста цветовая информация не несёт существенной смысловой нагрузки, а использование одного канала позволяет снизить вычислительную сложность модели. В процессе предобработки каждое изображение нормализуется к диапазону [0, 1] и приводится к размеру 28×28 пикселей.

1.2.3. Свёрточный слой: операция свёртки, ядро, stride и padding

Основным строительным блоком свёрточной нейронной сети выступает свёрточный слой, реализующий операцию свёртки. Как иллюстрирует рисунок 3, ядро (фильтр) небольшого размера последовательно перемещается по изображению и вычисляет поэлементное умножение значений пикселей на веса фильтра.

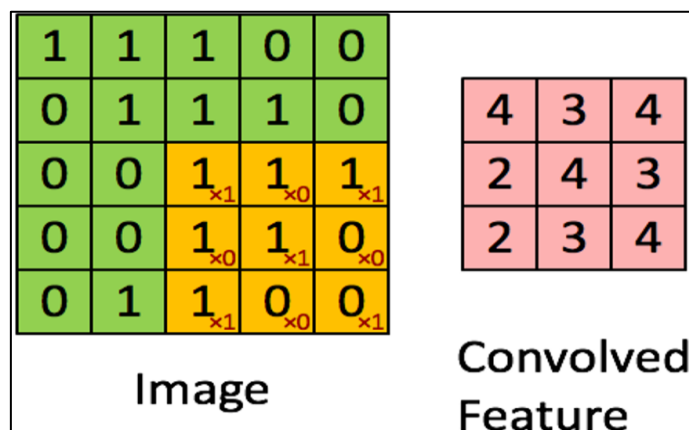


Рисунок 3 – Операция свёртки: пример применения ядра 3×3×1 к изображению 5×5 для получения свёрнутого признака 3×3×1

Важными параметрами свёрточного слоя являются размер ядра, шаг перемещения (stride) и способ обработки границ изображения (padding). На рисунке 4 демонстрируется перемещение ядра с шагом, равным двум, и пример нулевого заполнения (zero-padding).

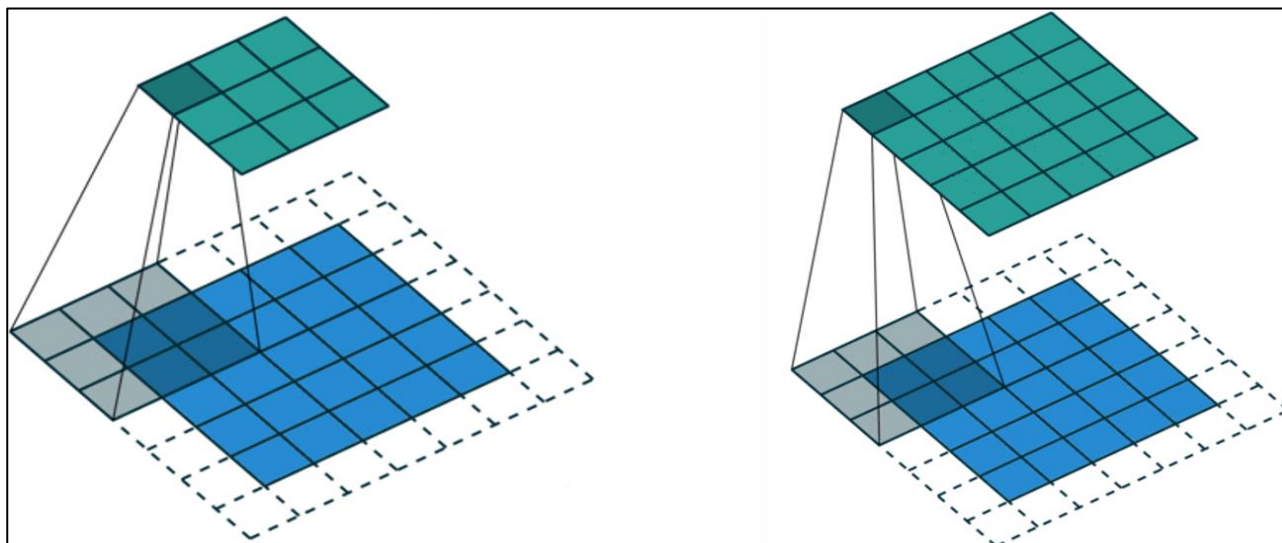


Рисунок 4 – Перемещение ядра свёртки с шагом, равным двум, и пример нулевого заполнения для сохранения размера выходного признака

1.2.4. Слой подвыборки (Pooling)

После свёрточных слоёв обычно следуют слои подвыборки, которые уменьшают пространственные размеры карт признаков и повышают устойчивость модели к небольшим сдвигам и искажениям. На рисунке 5 представлены пример объединения значений в окне и сравнение двух основных типов подвыборки – Max-Pooling и Average-Pooling.

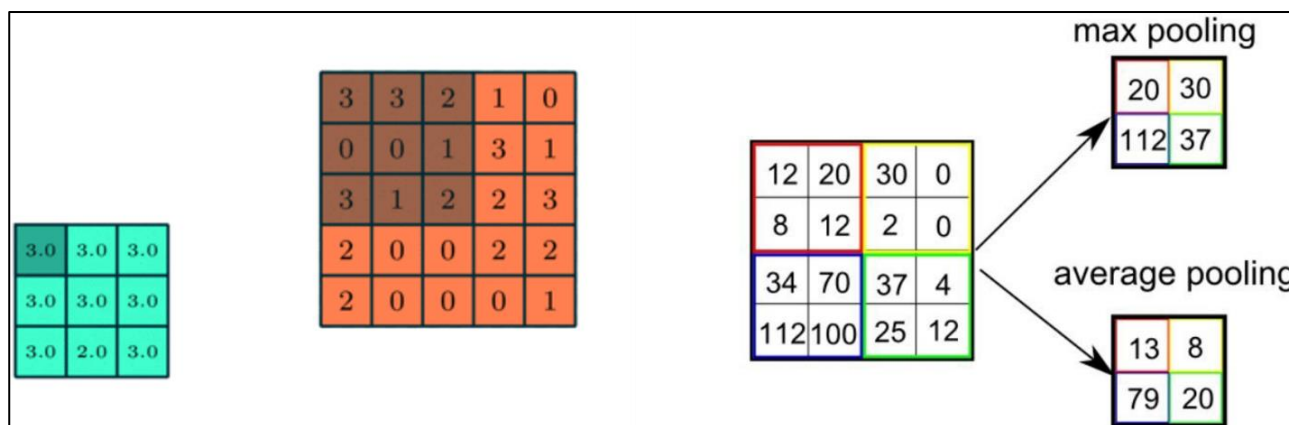


Рисунок 5 – Слой подвыборки: пример объединения 3×3 над свёрнутым признаком 5×5 и сравнение Max-Pooling с Average-Pooling

В разработанной модели применяется именно Max-Pooling с окном 2×2. Этот выбор обусловлен тем, что максимальное значение лучше сохраняет наиболее выраженные признаки символа, в то время как усреднение может сглаживать важные детали, особенно при работе с тонкими линиями и характерными элементами букв.

1.2.5. Полносвязный слой и классификация

После серии свёрточных и подвыборочных слоёв извлечённые признаки объединяются для принятия финального решения о классе символа. Как показано на рисунке 6, карты признаков сначала разворачиваются в одномерный вектор (Flatten), после чего поступают на полносвязный слой с функцией активации ReLU. Выход этого слоя подаётся на Softmax-слой, который преобразует значения в вероятности принадлежности к каждому из классов.

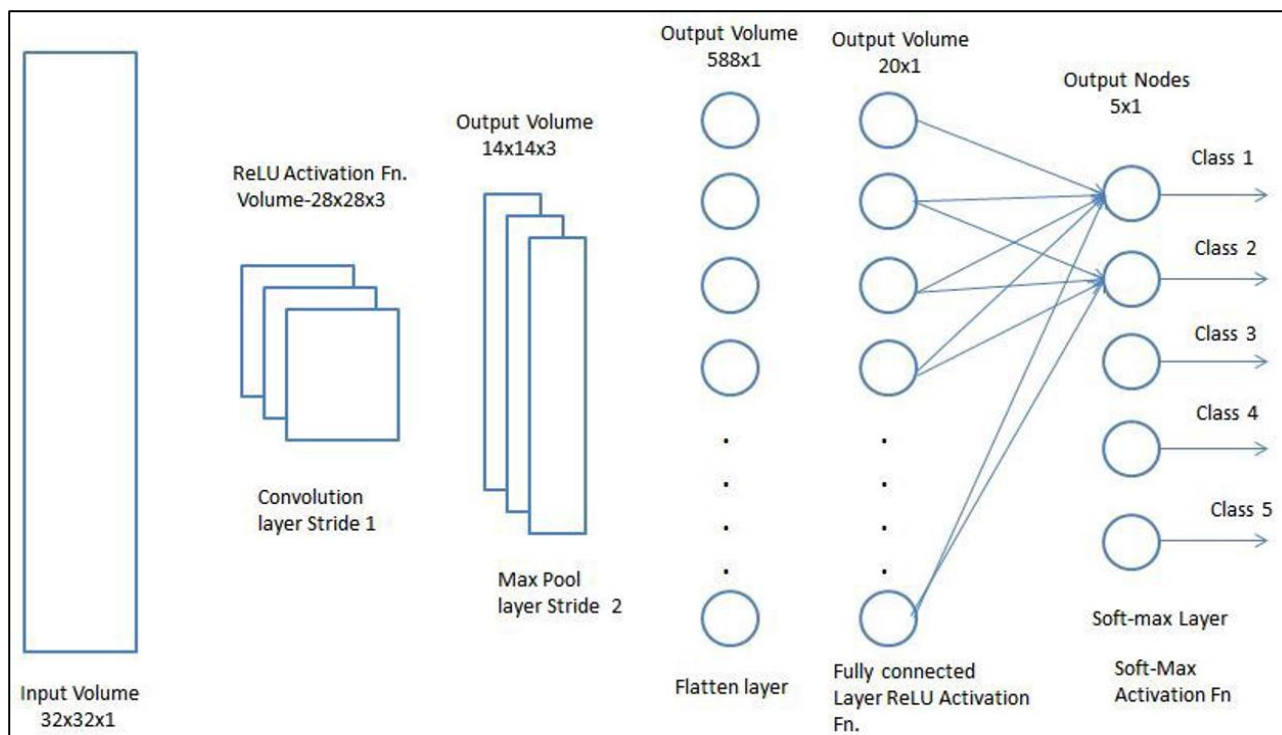


Рисунок 6 – Классификация через полносвязный слой Soft-max Layer

В модели приложения после Flatten используется полносвязный слой на 128 нейронов с функцией активации ReLU и Dropout для регуляризации. Выходной слой содержит 62 нейрона, соответствующих всем заглавным и строчным буквам английского алфавита вместе с цифрами. Функция Softmax обеспечивает интерпретируемый выход в виде вектора вероятностей, на основе которого в дальнейшем вычисляется базовая оценка качества написания.

1.3. Распознавание рукописного текста с помощью CNN: датасеты, вызовы и ограничения

Выбор набора данных для обучения свёрточной нейронной сети определяется практическими соображениями. Набор EMNIST ByClass был выбран потому, что он содержит все 62 класса – цифры от 0 до 9, а также заглавные и строчные буквы английского алфавита. Сравнительная характеристика поднаборов EMNIST представлена в табл. 1. Как видно из таблицы, именно EMNIST ByClass наиболее подходит для решаемой задачи, поскольку содержит оба регистра букв и цифры, а также обладает наибольшим объёмом данных.

Таблица 1

Сравнительная характеристика поднаборов EMNIST

Характеристика	EMNIST ByClass	EMNIST ByMerge	EMNIST Balanced	EMNIST Letters
Состав классов	0–9, A–Z, a–z (62 кл.)	Объединённые похожие символы (47 кл.)	То же, что ByMerge, сбалансировано (47 кл.)	A–Z (26 кл.)
Наличие заглавных и строчных букв	Оба регистра	Частично объединены	Частично объединены	Только заглавные
Наличие цифр	Да	Да	Да	Нет
Общее число изображений	814 255	814 255	131 600	145 600
Из них обучающая выборка	697 932	697 932	112 800	124 800
Из них тестовая выборка	116 323	116 323	18 800	20 800
Валидационная выборка	Отсутствует	Отсутствует	Отсутствует	Отсутствует

На рисунке 7 показан пример содержимого датасета EMNIST ByClass. Изображения имеют значительную вариативность начертания, что делает этот набор близким к реальному детскому почерку.

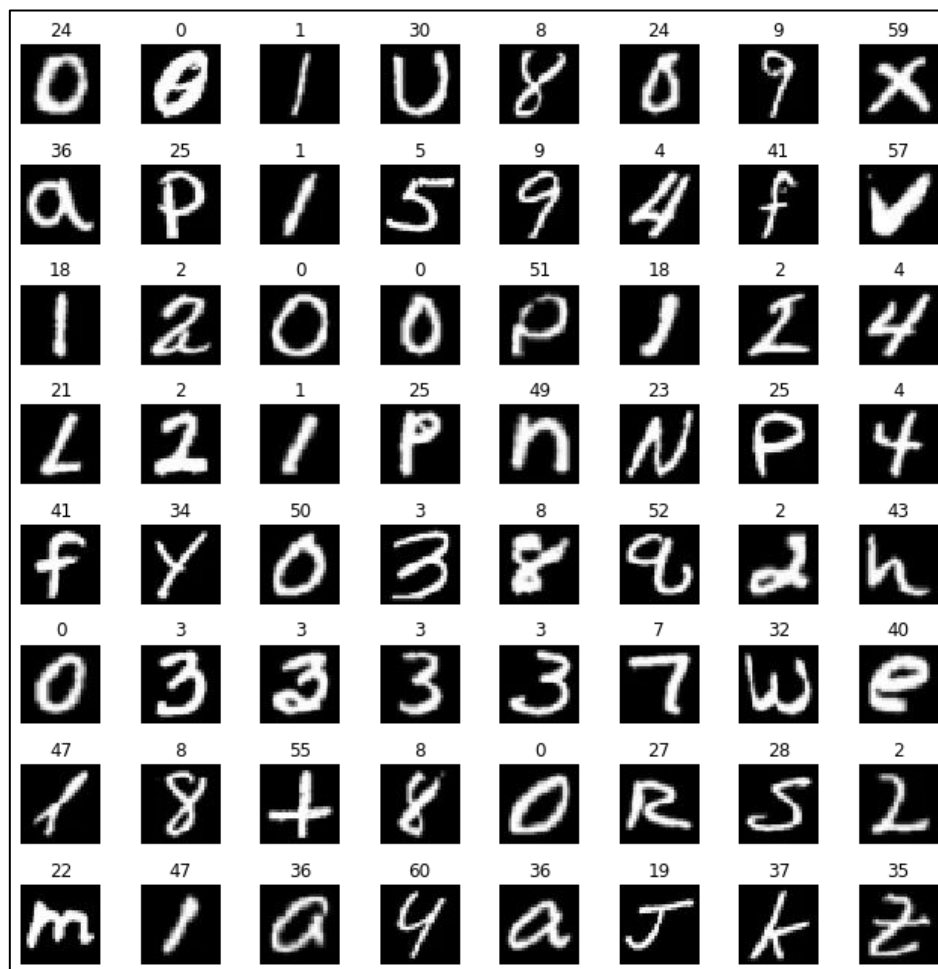


Рисунок 7 – Пример содержимого датасета EMNIST ByClass

В то же время работа с рукописным текстом, особенно детским, ставит перед моделью серьёзные вызовы. Детский почерк отличается высокой вариативностью размера, наклона, пропорций и положения символов относительно базовых линий. Кроме того, в написании часто присутствуют характерные элементы (точки над *i* и *j*, горизонтальные перекладки, замкнутые контуры). Эти особенности требуют от модели не только классификации, но и устойчивости к искажениям.

Стандартные технологии оптического распознавания символов недостаточно эффективны в образовательном контексте. Они хорошо справляются с идентификацией символа, однако не учитывают геометрические характеристики написания, которые критически важны для оценки качества почерка у ребёнка [8]. Именно это ограничение делает необходимым переход к гибридным решениям.

1.4. Гибридные подходы в задачах оценки качества рукописного ввода

Чисто классификационный подход на основе свёрточной нейронной сети не решает полностью задачу объективной оценки качества рукописного ввода. Нейронная сеть способна определить, какой символ был введён, однако не проверяет, насколько правильно он написан – соответствует ли положение относительно базовых линий, соблюдены ли пропорции и присутствуют ли необходимые характерные элементы.

Гибридные системы, сочетающие глубокое обучение с геометрическим анализом, позволяют преодолеть это ограничение. Свёрточная нейронная сеть отвечает за классификацию изображения, а модуль постобработки выполняет геометрический анализ и

корректирует результат. В разрабатываемом приложении геометрический анализ опирается на данные с клиентского холста – координаты границ символа, его высоту и положение относительно трёх направляющих линий. Эти параметры позволяют, например, отличить заглавную букву от строчной.

Кроме того, в модуле постобработки применяются эвристические правила для исправления типичных ошибок распознавания (анализ наличия точки над штрихом для различения *i* и *j*, поиск замкнутого контура для отличия *b* и *G*). Такие правила реализуются без дополнительных обучаемых параметров и сохраняют вычислительную эффективность системы.

Сочетание нейросетевой классификации с геометрическим и эвристическим анализом позволяет формировать не только факт распознавания, но и количественную оценку качества написания. В результате пользователь получает развёрнутую обратную связь, необходимую для формирования навыка аккуратного и разборчивого письма.

2. Разработка web-приложения и гибридной системы оценки качества рукописного ввода

2.1. Постановка задачи и функциональные требования

Разработка web-приложения велась в соответствии с техническим заданием, согласованным с Детским технопарком «Кванториум». Основное требование заказчика заключалось в создании инструмента, позволяющего детям младшего школьного возраста выполнять свободный рукописный ввод без жёстких шаблонов и получать мгновенную качественную обратную связь.

К функциональным требованиям относятся: интерактивный холст для свободного написания символов, распознавание введённого символа с помощью обученной свёрточной нейронной сети, автоматическая оценка качества написания на основе гибридной системы, работа с целыми словами в тестовом режиме, а также сохранение результатов попыток и статистики прогресса. Нефункциональные требования включают интуитивно понятный интерфейс и время обработки не более двух секунд.

2.2. Архитектура web-приложения

Архитектура web-приложения построена по классической клиент-серверной модели. Клиентская часть реализована с использованием HTML5, CSS и модульного JavaScript с применением Canvas API. На холсте размером 280×280 пикселей отображаются три направляющие линии. Серверная часть построена на основе фреймворка Flask и отвечает за маршрутизацию запросов, предобработку изображений, выполнение инференса и работу гибридной системы постобработки [3].

Взаимодействие между клиентом и сервером организовано через REST API. При нажатии кнопки проверки клиент отправляет на сервер предварительно обработанное и закодированное в Base64 изображение. Сервер выполняет полную предобработку, запускает модель вместе с модулем постобработки и возвращает результат в формате JSON.

2.3. Подготовка и аугментация набора данных

Основу обучающего набора данных составил набор EMNIST ByClass, содержащий 62 класса символов. Для повышения устойчивости модели к реальному детскому почерку исходный набор был дополнен синтетически сгенерированными образцами, выполненными шрифтом Caveat. Генерация образцов осуществлялась с помощью библиотек Pillow, OpenCV и NumPy: для каждого из 62 символов создавалось по 100 вариантов с случайными изменениями размера, поворотом в пределах ± 10 градусов и вариацией толщины штриха.

Аугментация позволила существенно улучшить обобщающую способность модели на данных, приближенных к реальному рукописному вводу. На рисунке 8 наглядно показано различие между «идеальным» образцом буквы А из EMNIST и «усреднённым» вариантом, сгенерированным шрифтом Caveat.



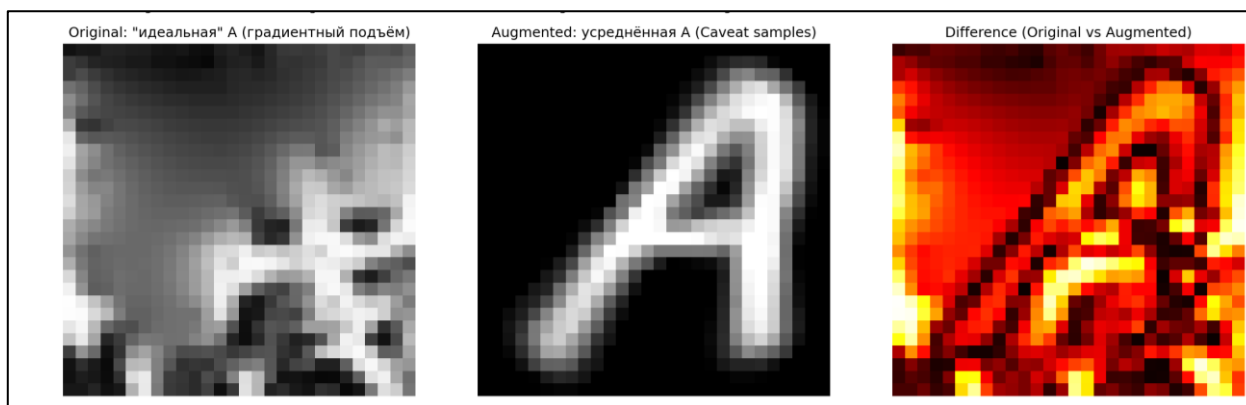


Рисунок 8 – Сравнение оригинального образца буквы А из EMNIST и аугментированного образца, сгенерированного шрифтом Caveat, вместе с визуализацией разницы между ними

2.4. Архитектура свёрточной нейронной сети, сравнение конфигураций и процесс обучения

В процессе разработки были рассмотрены две основные конфигурации свёрточной нейронной сети, различающиеся цветовой схемой фона: модель на чёрном фоне (ч/б) и модель на белом фоне (б/ч). Обе конфигурации строились по одной базовой схеме из двух свёрточных слоёв, Max-Pooling и полносвязного слоя.

На рисунке 9 представлены кривые обучения для обеих конфигураций при обучении на оригинальном наборе данных без аугментации. На рисунке 10 показаны те же кривые при обучении на расширенном аугментированном наборе.

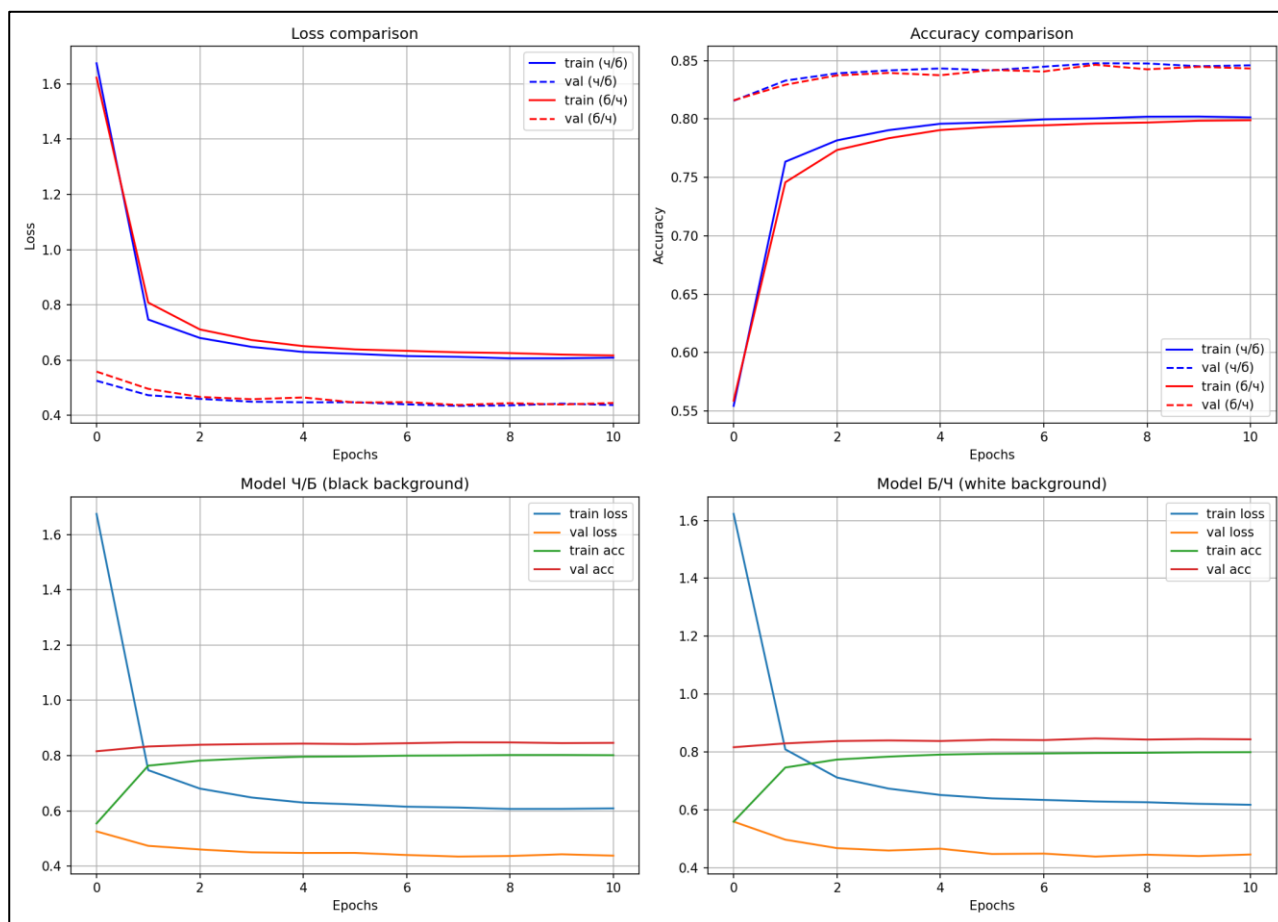


Рисунок 9 – Сравнение loss и accuracy при обучении моделей конфигураций ч/б (чёрный фон) и б/ч (белый фон) на оригинальном наборе данных без аугментации

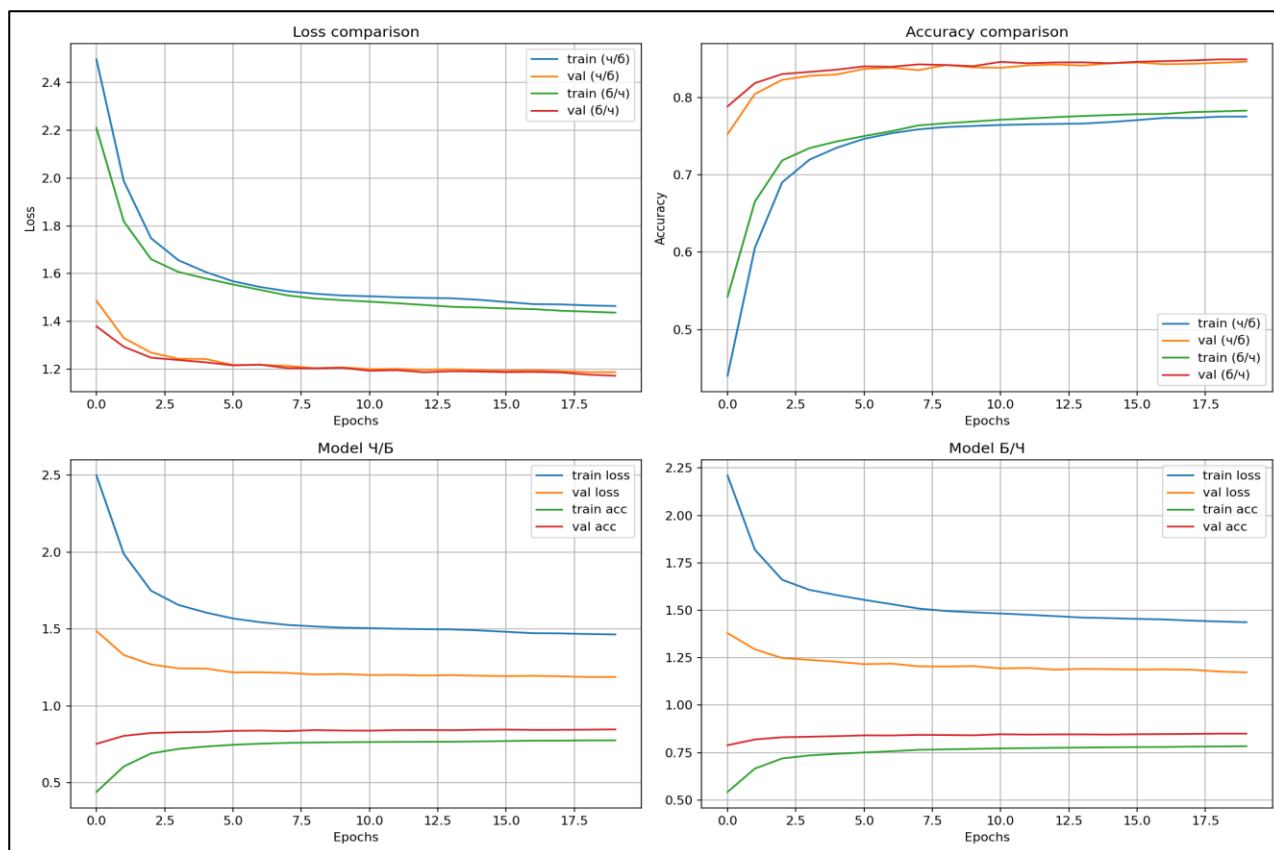


Рисунок 10 – Сравнение loss и ассигасу при обучении моделей конфигураций ч/б (чёрный фон) и б/ч (белый фон) на аугментированном наборе данных

Обучение проводилось с использованием трёх callback-функций: EarlyStopping для предотвращения переобучения, ModelCheckpoint для сохранения лучшей версии по валидационной метрике и DualTestCallback для одновременного мониторинга качества [2]. В конечном итоге была выбрана модель б/ч, обученная на аугментированном наборе (SELECTED_MODEL = 2). Структура данной модели приведена в табл. 2.

Таблица 2

Структура свёрточной нейронной сети (параметры слоёв)

Слой	Тип слоя	Фильтры / Нейроны	Размер ядра / окна	Активация	Прочие параметры
1	Входной	–	(28, 28, 1)	–	–
2	Свёрточный	32	3×3	ReLU	padding='same'
3	Подвыборочный	–	2×2	–	stride=2
4	Свёрточный	64	3×3	ReLU	padding='same'
5	Подвыборочный	–	2×2	–	stride=2
6	Преобразование	–	–	–	–
7	Полносвязный	128	–	ReLU	–
8	Регуляризация (Dropout)	–	–	–	rate=0.5
9	Выходной	62	–	Softmax	–

2.5. Гибридная система постобработки и оценки качества рукописного ввода

После получения предсказания от свёрточной нейронной сети результат проходит через модуль постобработки. Этот модуль выполняет две основные функции: корректировку предсказания на основе геометрических характеристик символа и расчёт итоговой количественной оценки качества написания.



Геометрический анализ использует данные, переданные с клиентской стороны: координаты минимальной и максимальной точек символа, полученные функцией `getBoundingBox`, а также высоту символа. На основании этих параметров определяется соответствие написания заглавному или строчному регистру. Дополнительно выполняются проверки на наличие точки над штрихом (для различения *i* и *j*) и наличие замкнутого контура внутри символа (для различения *b* и *G* с помощью функции `binary_fill_holes` из библиотеки `SciPy`).

На основе скорректированного предсказания вычисляется итоговая оценка `SCORE` в процентах. Далее применяется функция `get_verdict`, которая относит результат к одной из четырёх категорий: `Excellent`, `Good`, `Fair` или `Bad`.

2.6. Реализация клиентской части, механизмы хранения прогресса и аналитики

Клиентская часть приложения построена на модульной архитектуре `JavaScript`. Основные функциональные режимы – `Letter Mode`, `Test Mode`, `Best Attempts`, `Mistakes` и `Manual Test` – используют общую инфраструктуру холста и механизмы обмена данными с сервером через `Fetch API`.

Главное меню, представленное на рисунке 11, служит единой точкой входа. В режиме `Letter Mode` пользователь выполняет серию из десяти попыток для выбранного символа. Интерфейс этого режима показан на рисунке 12.

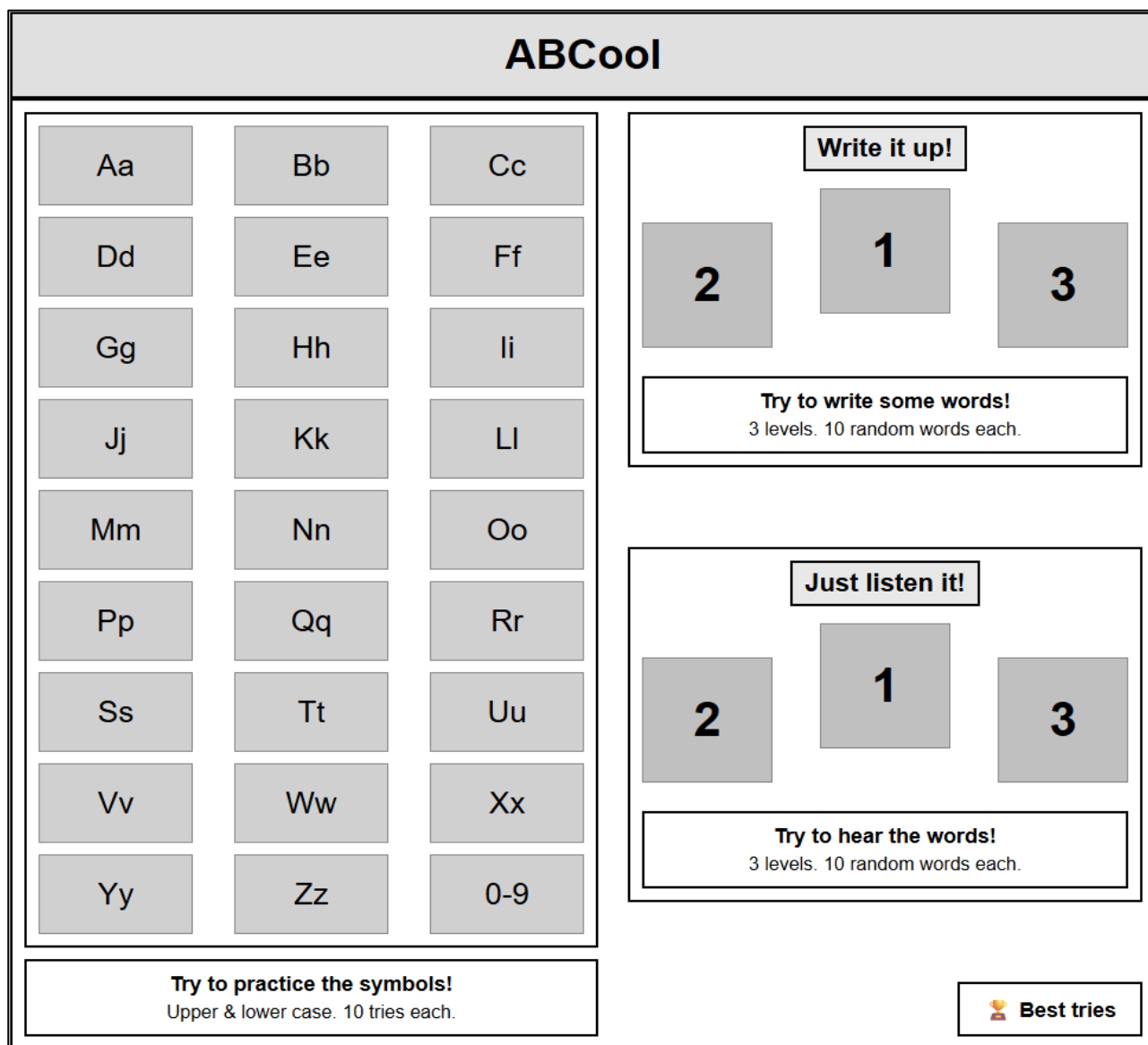


Рисунок 11 – Главное меню web-приложения
с выбором режимов тренировки и анализа прогресса

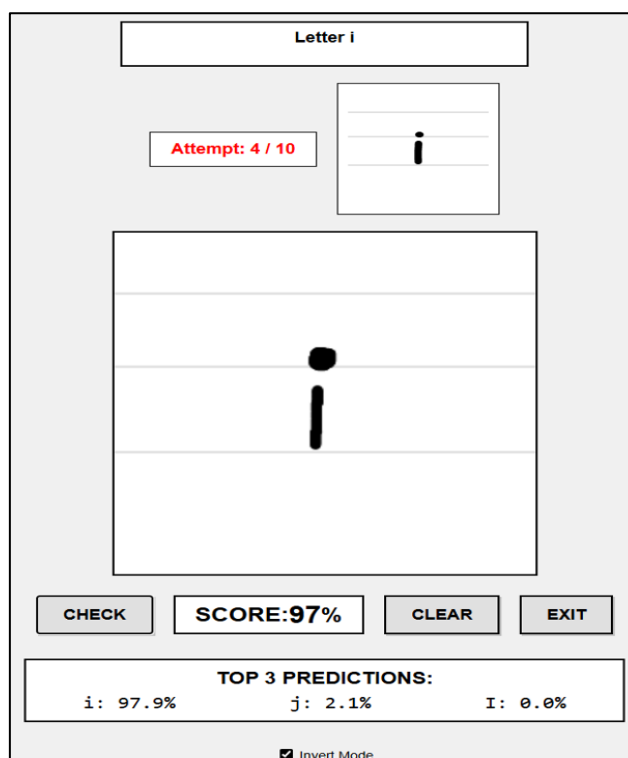


Рисунок 12 – Интерфейс режима Letter Mode с холстом, направляющими линиями и индикацией попыток

Режим Test Mode предназначен для проверки навыков написания целых слов. Функция segmentWord выполняет предварительную сегментацию слова на клиентской стороне. Интерфейс режима показан на рисунке 13.

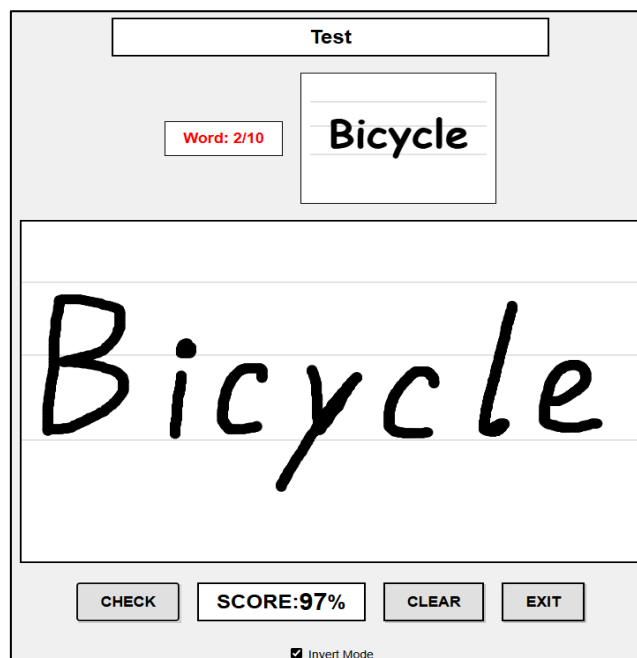


Рисунок 13 – Интерфейс режима Test Mode с холстом, адаптирующимся под длину слова, и результатами сегментации

Страница Best Attempts, изображённая на рисунке 14, формируется на сервере с помощью Jinja2 и отображает сводную статистику по всем символам. Страница Mistakes, показанная на рисунке 15, загружает данные в формате JSON и динамически формирует карточки с информацией об ошибочных словах.

[← Back to Home](#)
[Mistakes](#)
[Manual Test](#)

Best Attempts

Lowest Best Score

l 92%

h 98%

A 99%

a 99%

B 99%

b 99%

Lowest Average Score

G 37%

E 47%

D 53%

C 56%

b 62%

i 67%

All Symbols

Symbol	Best Attempts			Common Mistakes	Actions
A	99%	99%	99%	n: 1	
a	99%	99%	98%	no mistakes	
B	99%	99%	69%	g: 1	
b	99%	85%	66%	d: 1	
c	99%	98%	84%	o: 2	

Рисунок 14 – Страница Best Attempts с таблицей лучших попыток и статистикой по символам

Word Mistakes Gallery

[Delete All Mistakes](#)
[Reset All Stats](#)
[Back](#)

Word: Pig

SCORE: 83.33333333333333%

23.06.2026, 21:31:28

P → P (99%)

i → i (52%)

g → g (99%)

[Retest](#)
[Reset stats](#)
[Delete](#)

Word: House

SCORE: 79.2%

23.06.2026, 21:38:07

H → H (99%)

o → o (99%)

u → u (99%)

s → s (99%)

[Retest](#)
[Reset stats](#)
[Delete](#)

Word: Monster

SCORE: 28.285714285714285%

23.06.2026, 21:40:10

M → M (99%)

o → o (99%)

n → s (0%)

s → t (0%)

t → e (0%)

e → r (0%)

[Retest](#)
[Reset stats](#)
[Delete](#)

Word: Garden

SCORE: 82.5%

23.06.2026, 21:43:04

G → C (0%)

a → a (99%)

r → r (99%)

d → d (99%)

e → e (99%)

n → n (99%)

[Retest](#)
[Reset stats](#)
[Delete](#)

Рисунок 15 – Страница Mistakes с карточками ошибочных слов и возможностью повторной тренировки

Режим Manual Test, представленный на рисунке 16, позволяет пользователю самостоятельно вводить произвольное слово.



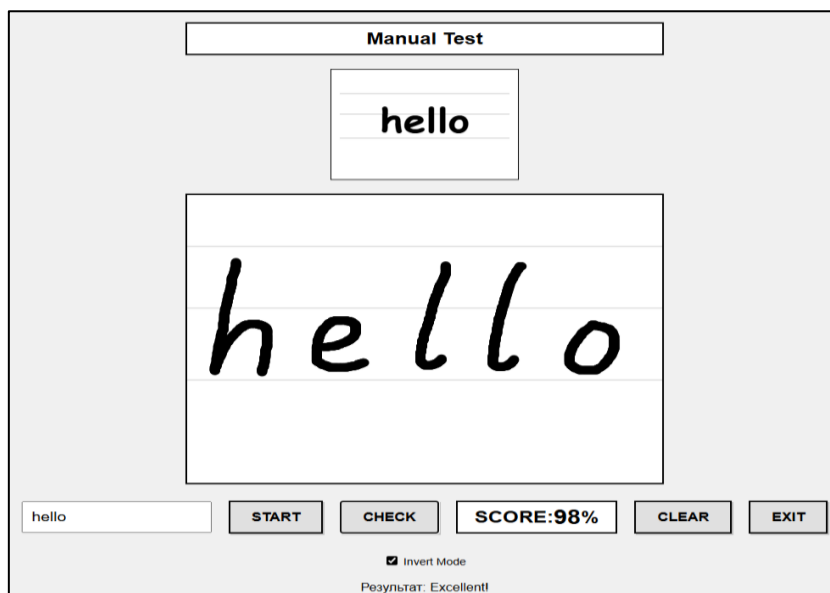


Рисунок 16 – Интерфейс режима Manual Test с динамически подстраивающимся холстом и мгновенным отображением результатов

Механизмы хранения прогресса реализованы на двух уровнях: временное хранение в localStorage браузера и постоянное хранение на сервере через модуль storage.py, который работает с двумя JSON-файлами – best_attempts.json и mistakes.json.

3. Результаты экспериментов и заключение

Экспериментальная часть работы была направлена на проверку эффективности разработанной гибридной системы оценки качества рукописного ввода. Основное внимание уделялось влиянию аугментации на точность модели.

В ходе обучения финальной модели на аугментированном наборе данных была достигнута точность 99,177 % на выборке, приближенной к реальному детскому почерку, при среднем времени инференса менее 500 миллисекунд. Для оценки вклада аугментации проведено сравнение двух версий модели: обученной на исходном наборе EMNIST (Original) и на расширенном наборе с добавлением синтетических образцов Caveat (Augmented). Результаты сравнения представлены на рисунке 17.

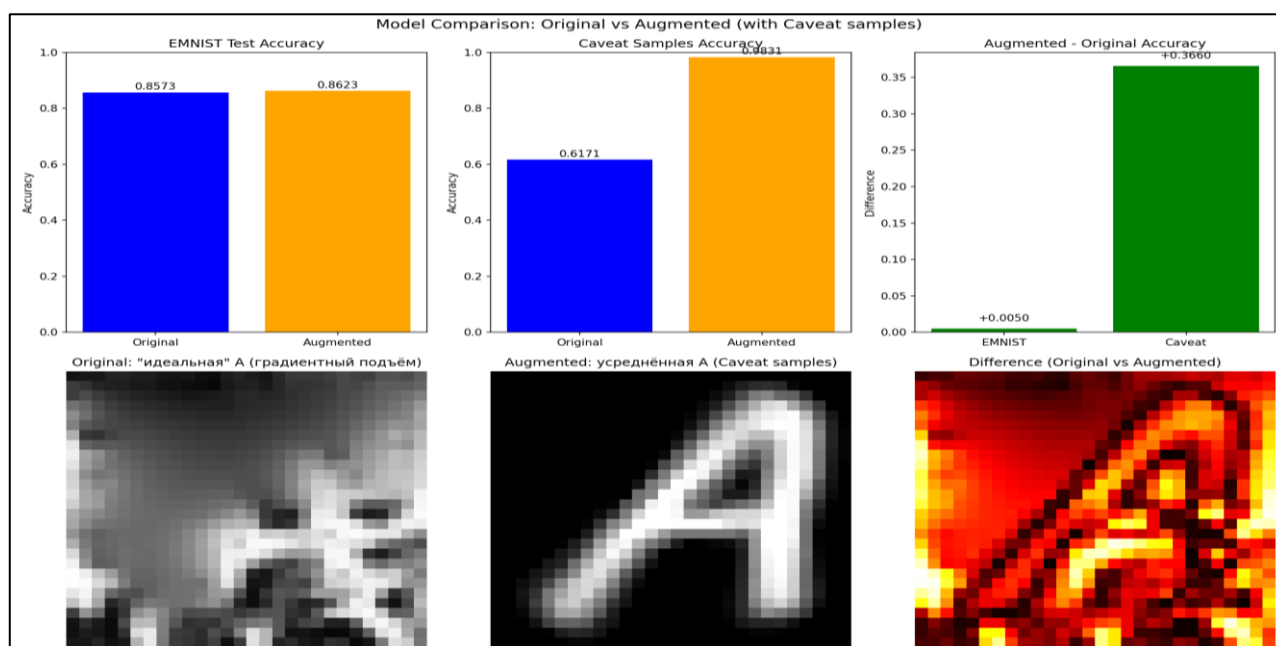


Рисунок 17 – Сравнение точности Original и Augmented моделей на тесте EMNIST и на выборке Caveat-образцов

Как видно из диаграмм, точность на стандартном тесте EMNIST практически не изменилась (0,8573 у Original против 0,8623 у Augmented). В то же время на выборке Caveat-образцов точность выросла с 0,6171 до 0,9831 (+0,366). Это подтверждает, что добавление синтетических образцов шрифта Caveat позволило модели значительно лучше обобщать знания на данных с вариативным начертанием символов.

Реализован механизм Dual Test Monitoring для одновременного отслеживания точности на двух независимых тестовых наборах – EMNIST и Caveat. Динамика изменения метрик представлена на рисунке 18.

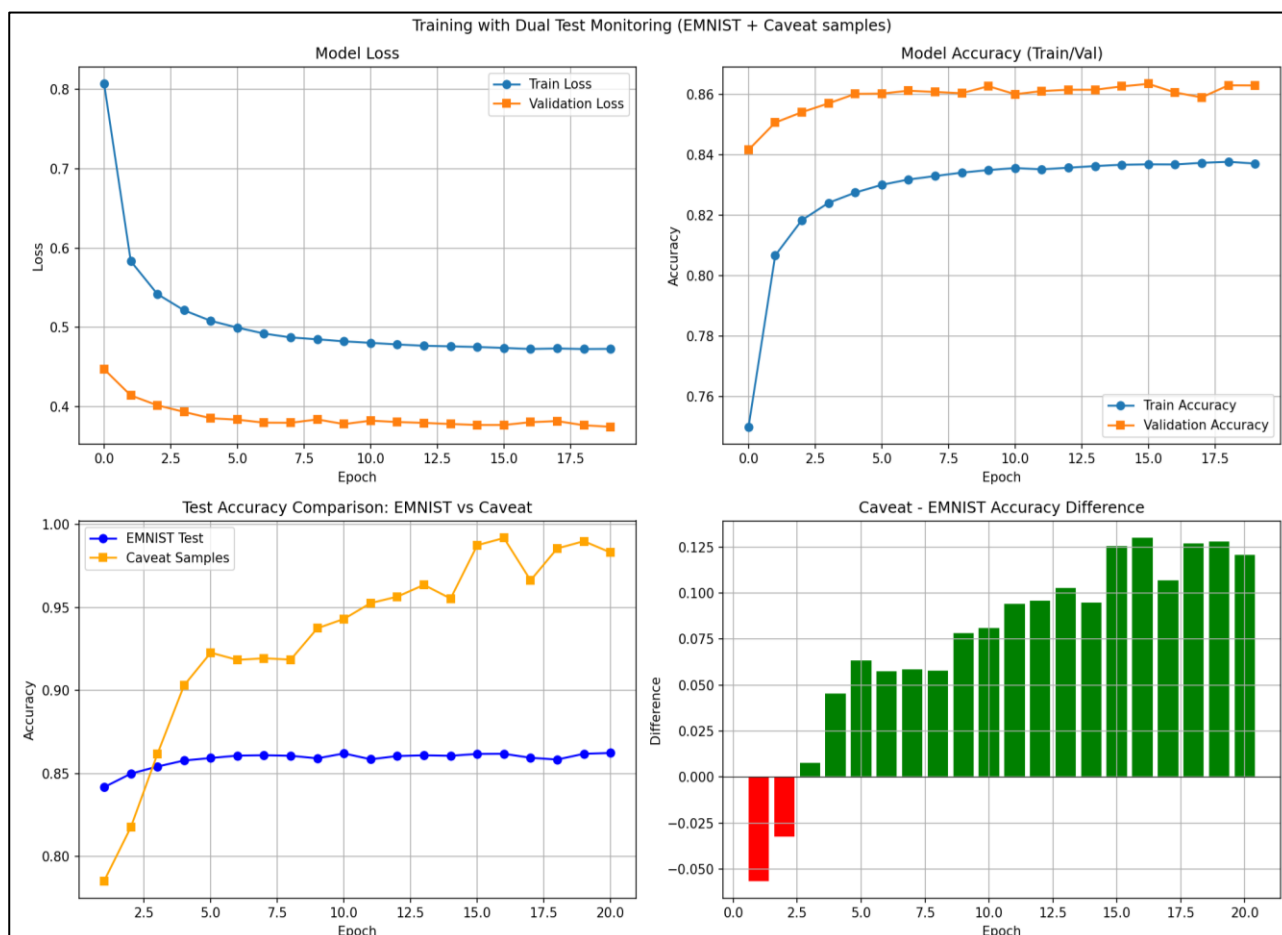


Рисунок 18 – Результаты Dual Test Monitoring: сравнение ассурасу на EMNIST и Caveat-образцах, а также разница между ними по эпохам

Точность на тесте EMNIST стабилизировалась около 0,86 уже после первых эпох. Точность на Caveat-образцах продолжала уверенно расти на протяжении всего процесса, достигнув значений выше 0,98. Разница между двумя метриками после начального периода отрицательных значений стала положительной и продолжала увеличиваться. Это подтверждает, что модель постепенно специализировалась на целевом домене (рукописный стиль Caveat), сохраняя при этом высокое качество на стандартных данных.

4. Заключение

В ходе работы все поставленные задачи были выполнены в полном объёме. Цель исследования – разработка web-приложения для обучения написанию букв английского алфавита, реализующего гибридную систему оценки качества рукописного ввода на основе свёрточной нейронной сети и геометрического анализа – достигнута.

Разработанное приложение обеспечивает объективную мгновенную оценку качества свободного рукописного ввода с учётом особенностей детского почерка. Гибридная система постобработки позволяет формировать развёрнутую обратную связь, включающую как количественную оценку, так и цветовую индикацию качества написания.

Приложение внедрено в образовательный процесс Детского технопарка «Кванториум» при МАОУ Гимназия № 3 г. Южно-Сахалинска. Оно сохраняет результаты попыток, лучшие образцы и статистику ошибок, предоставляя учащимся и педагогам инструменты для анализа индивидуального прогресса. После доработки пользовательского интерфейса приложение полностью готово к широкому использованию

Список литературы:

1. Волкова С. С. Основы теории нейронных сетей. – Вологда: Вологодский государственный университет, 2024.
2. Гаунов С. Р., Баймурадов У. Г., Ситников С. Ю. Машинное обучение на Python: использование библиотек Tensorflow и Scikit-Learn // Экономика и управление: проблемы, решения. – 2024. – Т. 8, № 12. – С. 72–81.
3. Косых Н. Е., Хомоненко А. Д., Куранова О. Н. Разработка WEB-приложения для анализа настроений текста с помощью фреймворка Flask и языка Python // Наукоемкие технологии в космических исследованиях Земли. – 2022. – Т. 14, № 1. – С. 45–52.
4. Крутько Д. А. Применение сверточных нейронных сетей в задаче распознавания рукописных цифр // Решетневские чтения: Материалы XXV Международной научно-практической конференции. – Красноярск: Сибирский государственный университет науки и технологий, 2021. – Ч. 2. – С. 376–377.
5. Моржов И. И., Сай С. В. Применение нейросетей на примере распознавания символов // Информационные технологии XXI века: Сборник научных трудов. – Хабаровск: Тихоокеанский государственный университет, 2020. – С. 214–220.
6. Самылкина Н. Н., Калинин И. А. Изучение работы нейронной сети на примере распознавания рукописных цифр с использованием языка программирования Python 3.8 // Информатика в школе. – 2021. – № 10. – С. 21–29.
7. Селедец И. Е., Борунов Д. И., Рычков В. А. Основы машинного обучения // Синергия Наук. – 2020. – № 44. – С. 407–414.
8. Rangasrinivasan S., Sumi S. S., Setlur S., Jayaraman B., Govindaraju V. From Scribbles to Text: A Novel Transformer-Based Recognition Model for Child Handwriting // Lecture Notes in Computer Science. – 2026. – Vol. 16023. – P. 115–131

