Программная система предварительной проверки самостоятельных работ по математическому анализу

Филимонова Елена Геннадиевна
кандидат экономических наук, доцент кафедры высшей математики и механики, Санкт-Петербургский государственный университет аэрокосмического приборостроения, elenaf3@yandex.ru,

Титов Вадим Евгеньевич
старший преподаватель, кафедра высшей математики и механики, Санкт-Петербургский государственный университет аэрокосмического приборостроения, vadimtitov460@gmail.com

В статье представлен прототип программного комплекса для автоматизированной обработки самостоятельных работ по теме «Неопределённый интеграл». Предлагаемый подход основан на применении стандартизированного бланка, содержащего отдельные области для указания ФИО и группы студента, записи условий заданий, выполнения вычислений и внесения окончательных ответов. Программа распознаёт идентификационные данные, определяет вариант работы и преобразует рукописные математические выражения в формализованную запись. Для обработки изображений предусмотрено использование предварительно обученных моделей, адаптированных на размеченных примерах заполнения бланков. Проверка итоговых ответов выполняется средствами символьных вычислений: распознанная первообразная дифференцируется, после чего полученное выражение сопоставляется с соответствующей подынтегральной функцией. Центральная часть листа с промежуточными вычислениями автоматически не анализируется, что позволяет снизить вычислительную нагрузку и вероятность ошибочной интерпретации неструктурированных записей. При недостаточной уверенности распознавания или невозможности однозначной проверки работа передаётся преподавателю. Применение комплекса позволяет ускорить первичную обработку самостоятельных работ и сократить количество повторяющихся операций при сохранении контроля со стороны преподавателя.

Ключевые слова: автоматизированная проверка, самостоятельная работа, рукописный ввод, распознавание математических выражений, неопределённый интеграл, первообразная, символьные вычисления, стандартизированный бланк, нейронные сети, образовательные технологии.

Введение

Проверка самостоятельных работ по математическим дисциплинам является одной из наиболее трудоёмких составляющих учебного процесса. При большом количестве студентов преподавателю приходится обрабатывать значительный объём однотипных материалов: устанавливать принадлежность работы конкретному обучающемуся, определять выполненный вариант задания, находить итоговые ответы, сопоставлять их с эталонными значениями и фиксировать результаты проверки. Ручное выполнение этих операций требует существенных временных затрат и может сопровождаться техническими ошибками, возникающими при идентификации работ, чтении рукописных записей и переносе оценок в ведомость.

Один из подходов к сокращению нагрузки на преподавателя заключается в использовании стандартизированного бланка самостоятельной работы, содержащего заранее определённые области для записи исходных данных, решения и итогового ответа. В ранее разработанном программном комплексе была предложена обработка подобных бланков применительно к самостоятельным работам по линейной алгебре [1]. Система объединяла выделение значимых областей документа, распознавание рукописных данных и математическую проверку полученных результатов. Дальнейшее развитие программного комплекса было связано с добавлением средств учёта проверенных работ, позволяющих сохранять сведения о студенте, учебной группе, номере работы и полученных баллах в общей таблице результатов [2].

Вместе с тем перенос данного подхода на задания по математическому анализу требует отдельного исследования. В частности, при проверке заданий на нахождение неопределённых интегралов система должна работать не только с текстовыми данными, но и с рукописными математическими выражениями, содержащими дроби, степени, скобки, тригонометрические функции и другие элементы математической нотации. Дополнительная сложность заключается в том, что один и тот же правильный ответ может быть записан в нескольких эквивалентных формах. Для неопределённых интегралов также необходимо учитывать наличие произвольной постоянной и возможность равносильных преобразований полученного выражения [3-5].

Предлагаемая система ориентирована на обработку стандартизированного листа самостоятельной работы, в котором выделяются три функциональные области. Первая область содержит идентификационные сведения — фамилию, имя и отчество студента, а также обозначение учебной группы. Во второй области записываются условия предложенных заданий. Третья область предназначена для итоговых ответов. Наличие фиксированной разметки позволяет выполнять геометрическое выделение указанных фрагментов изображения и применять к каждому из них специализированные методы распознавания [7].

Принципиальным ограничением разрабатываемой системы является отказ от автоматического анализа центральной части листа, содержащей последовательность рассуждений и промежуточные математические преобразования. Ход решения сохраняется на исходном изображении и при необходимости может быть просмотрен преподавателем, однако он не участвует в автоматическом формировании результата проверки. Такое ограничение обусловлено высокой вариативностью рукописного решения, отсутствием единого способа расположения промежуточных вычислений и сложностью надёжного установления логических связей между отдельными фрагментами записи [6, 9].

Таким образом, разрабатываемая программа должна выполнять предварительную проверку самостоятельной работы по её формализованным элементам. После загрузки фотографии или сканированного изображения листа система определяет границы бланка, выделяет области с идентификационными данными, условиями заданий и ответами, распознаёт их содержимое и сопоставляет полученные ответы с эталонными результатами. При недостаточной уверенности распознавания или невозможности однозначного сравнения работа должна передаваться преподавателю для ручной проверки [8].

Целью настоящего исследования является разработка программного средства для автоматизированной обработки стандартизированных листов самостоятельных работ по теме «Неопределённый интеграл», обеспечивающего распознавание сведений о студенте, идентификацию условий заданий и проверку итоговых ответов без анализа промежуточного хода решения. Практическая значимость разработки заключается в сокращении количества повторяющихся действий преподавателя, ускорении первичной обработки работ и формировании структурированных результатов контроля при сохранении возможности ручной проверки спорных случаев.

Требования к входным данным, структура бланка и адаптация модели распознавания

Надёжность автоматизированной проверки зависит от структуры документа и качества его изображения. При свободном расположении рукописных записей условия, вычисления и ответы могут пересекаться или находиться в разных частях листа. Поэтому в системе используется стандартизированный бланк с заранее определёнными функциональными областями.

Аналогичный подход применялся в разработанных ранее системах проверки работ по линейной алгебре [1, 2].

Пример бланка представлен на рисунке 1. Он содержит номер самостоятельной работы, поля для указания группы и ФИО студента, область условий, свободное пространство для вычислений и строку для окончательных ответов. Автоматически обрабатываются три области: идентификационные данные, условия заданий и ответы. Центральная часть с ходом решения сохраняется для просмотра преподавателем, но программой не анализируется. Такое ограничение позволяет уменьшить вычислительную нагрузку и вероятность ошибок, поскольку распознавание свободно расположенных промежуточных записей требует значительно более сложных алгоритмов и большого объёма обучающих данных. В результате система сосредоточена на обработке наиболее структурированных и значимых элементов работы [1].

Поле группы распознаётся как ограниченная последовательность букв и цифр, а полученное ФИО сопоставляется со списком студентов. Область условий используется для определения варианта и поиска соответствующих эталонных данных. Поскольку математические выражения имеют двумерную структуру, результат их распознавания преобразуется в формализованную запись, пригодную для символьной обработки.

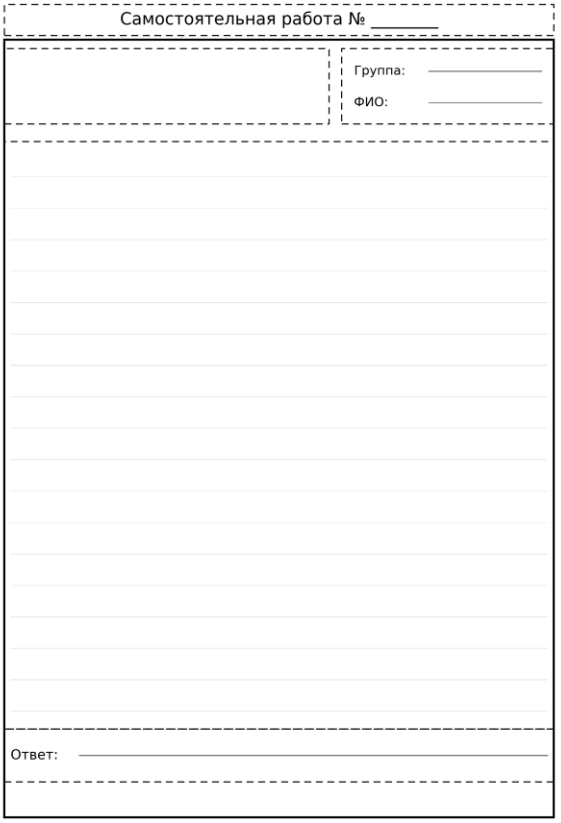


Рисунок 1 — Стандартизированный бланк самостоятельной работы

Для распознавания применяется не нейронная сеть, обучаемая с нуля, а предварительно обученная модель, адаптированная к используемому бланку. Для её дообучения формируется размеченная выборка. Из выровненных изображений работ извлекаются фрагменты с ФИО, группой, условиями и ответами, после чего каждому фрагменту вручную ставится в соответствие правильная текстовая или математическая запись. Формулы представляются в линейной нотации, близкой к LaTeX.

Выборка включает образцы почерка разных студентов и разделяется на обучающую, валидационную и тестовую части. Работы одного студента помещаются только в одну часть выборки, что позволяет проверить способность модели распознавать ранее неизвестный почерк. Для увеличения количества примеров применяются небольшие повороты, масштабирование, изменение контрастности, добавление шума, размытия и перспективных искажений.

Для обработки документа используются две специализированные модели. Первая распознаёт ФИО и обозначение группы, вторая преобразует изображения математических выражений в последовательность формальных символов. Словарь математической модели ограничивается элементами, встречающимися в заданиях на интегрирование: цифрами, переменными, арифметическими операциями, дробями, степенями, скобками, корнями, тригонометрическими функциями и постоянной C.

Результаты нейросетевого распознавания дополнительно проверяются формальными правилами. ФИО сопоставляется со списком студентов, группа — с известными обозначениями, а формула проверяется на правильность расположения скобок, операций, степеней и дробей. Каждый результат сопровождается оценкой уверенности. При недостаточной уверенности преподавателю показываются исходный фрагмент и предполагаемая расшифровка с возможностью ручного исправления.

В качестве входных данных используется фотография или сканированное изображение полного листа без сильного размытия, теней и бликов. После загрузки система обнаруживает контур документа, исправляет перспективу, нормализует яркость и выделяет необходимые области. Если границы бланка или содержимое поля невозможно определить достаточно надёжно, автоматическая обработка прекращается, и работа передаётся преподавателю.

Принципы организации символьных и численных вычислений

После загрузки изображения программа выполняет предварительную обработку документа, выделяет заданные области и распознаёт содержащиеся в них сведения. Результаты выводятся в интерфейсе, представленном на рисунке 2.

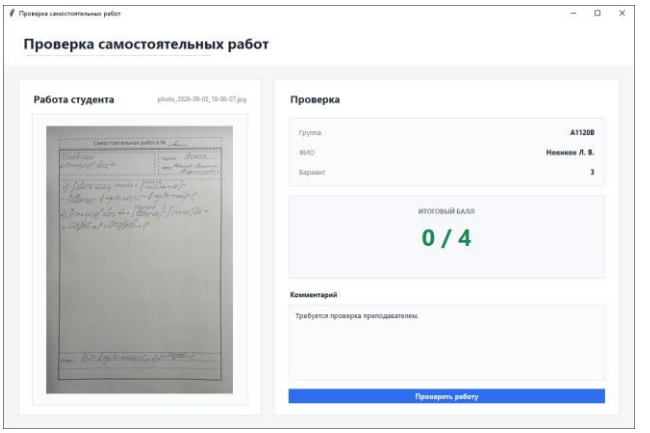


Рисунок 2 — Результат проверки самостоятельной работы

В левой части интерфейса отображается исходная фотография работы студента. Это позволяет преподавателю сопоставить автоматически полученные сведения с оригиналом и при необходимости просмотреть ход решения. Кнопки в нижней части окна предназначены для загрузки нового изображения и очистки рабочей области.

В правой части выводятся результаты распознавания. В представленном примере программа определила учебную группу A1120B, ФИО студента — Новиков Л. В., а также третий вариант самостоятельной работы. Распознанные данные могут быть дополнительно проверены по списку студентов и базе вариантов заданий.

Максимальная оценка за работу составляет четыре балла: по два балла за каждый из двух интегралов. После распознавания итоговых ответов программа преобразует их в символьную форму и проверяет посредством дифференцирования полученных первообразных. Если производная ответа не совпадает с соответствующей подынтегральной функцией, ответ считается неверным.

В рассматриваемой работе оба итоговых ответа не прошли математическую проверку. Поэтому за каждое задание было начислено ноль баллов, а общий предварительный результат составил 0 из 4 баллов. Ход решения при выставлении этого результата автоматически не анализировался.

Поскольку программа проверяет только итоговые ответы и не оценивает промежуточные преобразования, в поле комментария выводится сообщение «Требуется проверка преподавателем». Преподаватель может просмотреть исходное изображение, убедиться в правильности распознавания формул и после этого подтвердить или изменить результат.

Таким образом, программный комплекс не заменяет преподавателя полностью, а выполняет первичную обработку работы, распознаёт основные данные, проверяет итоговые ответы и выделяет случаи, требующие дополнительного контроля. Такой подход сокращает количество повторяющихся операций и одновременно сохраняет возможность принятия окончательного решения преподавателем.

Заключение

В работе представлен прототип программного комплекса для автоматизированной проверки самостоятельных работ по теме «Неопределённый интеграл». Основой системы является стандартизированный бланк, на котором выделены области с ФИО и группой студента, условиями заданий и итоговыми ответами. Центральная часть листа с промежуточными вычислениями автоматически не анализируется и остаётся доступной преподавателю для просмотра.

Для обработки рукописных данных предлагается использовать предварительно обученные модели, дополнительно адаптированные на размеченных примерах заполнения бланков. Распознанные ФИО сопоставляются со списком студентов, условия — с базой вариантов, а ответы преобразуются в символьную форму. Правильность первообразной определяется посредством её дифференцирования и сравнения полученной производной с подынтегральной функцией.

Практическое применение комплекса позволит сократить количество повторяющихся операций при обработке самостоятельных работ и ускорить получение предварительных результатов. Дальнейшее развитие системы связано с расширением обучающей выборки, повышением точности распознавания математических выражений, добавлением новых типов заданий и проведением экспериментальной оценки на работах разных студентов.

Литература

1. Аман Е. Э., Титов В. Е. Программный комплекс с элементами искусственного интеллекта для автоматизированной проверки самостоятельных работ по линейной алгебре // Инновации и инвестиции. 2026. № 4. С. 507–509.
2. Аман Е. Э., Пушкина В. П., Титов В. Е. Система проверки и учета работ по линейной алгебре // Инновации и инвестиции. 2026. № 6. С. 445–447.
3. Беляев С. А., Лапцевич Д. А. Подходы к созданию калькулятора с рукописным вводом // Программные продукты, системы и алгоритмы. 2018. № 1. С. 33–43. DOI: 10.15827/2311-6749.18.1.6.
4. Chan K.-F., Yeung D.-Y. Mathematical Expression Recognition: A Survey // International Journal on Document Analysis and Recognition. 2000. Vol. 3. P. 3–15. DOI: 10.1007/PL00013549.
5. Zhelezniakov D., Zaytsev V., Radyvonenko O. Online Handwritten Mathematical Expression Recognition and Applications: A Survey // IEEE Access. 2021. Vol. 9. P. 38352–38373. DOI: 10.1109/ACCESS.2021.3063413.
6. Graves A., Liwicki M., Fernández S., Bertolami R., Bunke H., Schmidhuber J. A Novel Connectionist System for Unconstrained Handwriting Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009. Vol. 31, no. 5. P. 855–868. DOI: 10.1109/TPAMI.2008.137.
7. Li M., Lv T., Chen J., Cui L., Lu Y., Florencio D., Zhang C., Li Z., Wei F. TrOCR: Transformer-Based Optical Character Recognition with Pre-trained Models // Proceedings of the AAAI Conference on Artificial Intelligence. 2023. Vol. 37, no. 11. P. 13094–13102. DOI: 10.1609/aaai.v37i11.26538.
8. Zhang J., Du J., Zhang S., Liu D., Hu Y., Hu J., Wei S., Dai L. Watch, Attend and Parse: An End-to-End Neural Network Based Approach to Handwritten Mathematical Expression Recognition // Pattern Recognition. 2017. Vol. 71. P. 196–206. DOI: 10.1016/j.patcog.2017.06.017.
9. Meurer A., Smith C. P., Paprocki M., Čertík O., Kirpichev S. B., Rocklin M., Kumar A., Ivanov S., Moore J. K., Singh S., Rathnayake T., Vig S., Granger B. E., Muller R. P., Bonazzi F., Gupta H., Vats S., Johansson F., Pedregosa F., Curry M. J., Terrel A. R., Roučka Š., Saboo A., Fernando I., Kulal S., Cimrman R., Scopatz A. SymPy: Symbolic Computing in Python // PeerJ Computer Science. 2017. Vol. 3. Article e103. DOI: 10.7717/peerj-cs.103.

Software System for Preliminary Assessment of Independent Work in Mathematical Analysis

Filimonova E.G., Titov V.E.
Saint Petersburg State University of Aerospace Instrumentation

JEL classification: C10, C50, C60, C61, C80, C87, C90

The article presents a prototype software system for the automated processing of student assignments on indefinite integrals. The proposed approach is based on a standardized form containing separate areas for the student's full name and group identifier, problem statements, calculations, and final answers. The system recognizes identification data, determines the assignment variant, and converts handwritten mathematical expressions into a formal representation. Pre-trained models adapted using annotated examples of completed forms are employed for image recognition. Final answers are verified by means of symbolic computation: the recognized antiderivative is differentiated, and the resulting expression is compared with the corresponding integrand. The central area containing intermediate calculations is not analyzed automatically. This limitation reduces computational requirements and lowers the risk of incorrectly interpreting unstructured handwritten notes. When recognition confidence is insufficient or an answer cannot be verified unambiguously, the assignment is referred to the instructor for manual review. The proposed system accelerates the preliminary processing of student assignments and reduces repetitive checking operations while preserving the instructor's control over the final assessment.

Keywords: automated assessment, student assignments, handwriting recognition, mathematical expression recognition, indefinite integral, antiderivative, symbolic computation, standardized form, neural networks, educational technologies.

References

1. Aman E. E., Titov V. E. Software Package with Elements of Artificial Intelligence for Automated Assessment of Independent Work in Linear Algebra // Innovations and Investments. 2026. No. 4. P. 507–509.
2. Aman E. E., Pushkina V. P., Titov V. E. A System for Checking and Recording Assignments in Linear Algebra // Innovations and Investments. 2026. No. 6. P. 445–447.
3. Belyaev S. A., Laptsevich D. A. Approaches to Creating a Calculator with Handwriting Input // Software Journal: Theory and Applications. 2018. No. 1. P. 33–43. DOI: 10.15827/2311-6749.18.1.6.
4. Chan K.-F., Yeung D.-Y. Mathematical Expression Recognition: A Survey // International Journal on Document Analysis and Recognition. 2000. Vol. 3. P. 3–15. DOI: 10.1007/PL00013549.
5. Zhelezniakov D., Zaytsev V., Radyvonenko O. Online Handwritten Mathematical Expression Recognition and Applications: A Survey // IEEE Access. 2021. Vol. 9. P. 38352–38373. DOI: 10.1109/ACCESS.2021.3063413.
6. Graves A., Liwicki M., Fernández S., Bertolami R., Bunke H., Schmidhuber J. A Novel Connectionist System for Unconstrained Handwriting Recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009. Vol. 31, No. 5. P. 855–868. DOI: 10.1109/TPAMI.2008.137.
7. Li M., Lv T., Chen J., Cui L., Lu Y., Florencio D., Zhang C., Li Z., Wei F. TrOCR: Transformer-Based Optical Character Recognition with Pre-trained Models // Proceedings of the AAAI Conference on Artificial Intelligence. 2023. Vol. 37, No. 11. P. 13094–13102. DOI: 10.1609/aaai.v37i11.26538.
8. Zhang J., Du J., Zhang S., Liu D., Hu Y., Hu J., Wei S., Dai L. Watch, Attend and Parse: An End-to-End Neural Network Based Approach to Handwritten Mathematical Expression Recognition // Pattern Recognition. 2017. Vol. 71. P. 196–206. DOI: 10.1016/j.patcog.2017.06.017.
9. Meurer A., Smith C. P., Paprocki M., Čertík O., Kirpichev S. B., Rocklin M., Kumar A., Ivanov S., Moore J. K., Singh S., Rathnayake T., Vig S., Granger B. E., Muller R. P., Bonazzi F., Gupta H., Vats S., Johansson F., Pedregosa F., Curry M. J., Terrel A. R., Roučka Š., Saboo A., Fernando I., Kulal S., Cimrman R., Scopatz A. SymPy: Symbolic Computing in Python // PeerJ Computer Science. 2017. Vol. 3. Article e103. DOI: 10.7717/peerj-cs.103.