

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН

Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И. Сатпаева»

Горно-металлургический институт имени О.А. Байконурова

Кафедра «Маркшейдерского дела и геодезии»

ДОПУЩЕН К ЗАЩИТЕ
НАО «КазНТУ им.К.И.Сатпаева»
Горно-металлургический институт
им. О.А. Байконурова

ДОПУЩЕН К ЗАЩИТЕ
Заведующий кафедрой
«Маркшейдерского дела и геодезии»
к.т.н., асс. профессор
Мейрамбек Г.
«13» 06 2025 г.

ДИПЛОМНАЯ РАБОТА

На тему: «Автоматизация процессов оцифровки и обработки геопространственных данных с
использованием современных ГИС-технологий»

6B07304 – Геопространственная цифровая инженерия

Выполнил

Мирсидиков Ш.Х.

Рецензент
МОК

к.т.н., асс. профессор

Омиржанова Ж.Т.

«13» 06 2025 г.

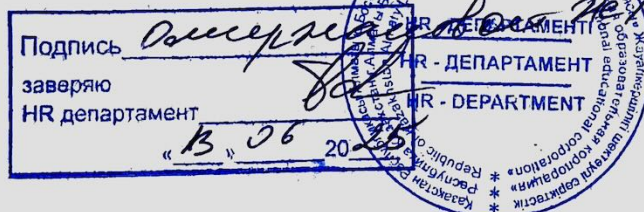
Научный руководитель

КазНТУ им.К.И.Сатпаева

PhD, ст. преподаватель

Камза А.Т.

«13» 06 2025 г.



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН

Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И. Сатпаева»

Горно-металлургический институт имени О.А. Байконурова

Кафедра «Маркшейдерского дела и геодезии»

6B07304 – Геопространственная цифровая инженерия

УТВЕРЖДАЮ

Заведующий кафедрой
«Маркшейдерского дела и геодезии»
к.т.н., доцент, профессор
Мейрамбек Г.

«10» 106 20 г.

ЗАДАНИЕ

на выполнение дипломной работы

Обучающемуся Мирсидиков Шахнур Хожасбарович

Тема: Автоматизация процессов оцифровки и обработки геопространственных данных с использованием современных ГИС-технологий

Утверждена приказом Проректора по академическим вопросам №26-П/Ө от 29.01.2025г.

Срок сдачи законченной работы «05» июня 2025г.

Исходные данные к дипломной работе: *Материал, накопленный во время производственной практики, данные лекций, сканы геологических карт*

Краткое содержание дипломной работы:

- а) анализ теоретических основ цифровизации геопространственных данных и роли ГИС-технологий;
- б) обзор и сравнение инструментов автоматической векторизации;
- в) обоснование применения оцифрованных карт для нужд кадастра и инженерных изысканий;
- г) вопросы безопасности жизнедеятельности и охраны труда; расчет экономической эффективности внедрения автоматизации.

Перечень графического материала: *представлены 20 слайдов презентации работы*

Рекомендуемая основная литература:

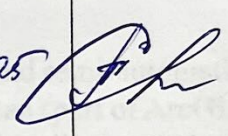
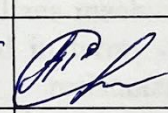
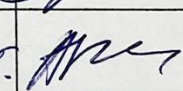
1. Жеңісова Н.Е., Тұрымтаев Ж.Б., Қазықанова А.Қ., Токтаров Ж.А. Технология создания цифровой картографической основы для информационно-оценочных карт территории Республики Казахстан: диссертация / КазНУ им. аль-Фараби, 2023. – 150 с.
2. Скворцова Р.Н., Согорова И.Н. Геодезия, кадастр с основами геоинформатики: учебник для вузов / Р.Н. Скворцова, И.Н. Согорова; под ред. Р.И. Золотовой. — Алматы: Казахский техн. ун-т, 2021. — 320 с
3. Зулпыхаров К.Б. Геоинформационные системы: назначение, функции и методы векторного преобразования данных: учебное пособие / КазНУ им. аль-Фараби, 2024—200 с

ГРАФИК
подготовки дипломной работы (проекта)

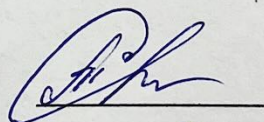
Наименование разделов, перечень разрабатываемых вопросов	Сроки представления научному руководителю	Примечание
Теоретические основы автоматизации оцифровки геопространственных данных и их роль в кадастровой деятельности	15.02.2025г.	нет
Обзор и анализ технологий автоматической оцифровки	17.04.2025г.	нет
Разработка и реализация метода автоматической оцифровки на примере геологических карт	14.05.2025г.	нет

Подписи

консультантов и норм контролера на законченную дипломную работу (проект)с
указанием относящихся к ним разделов работы (проекта)

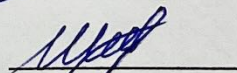
Наименования разделов	Консультанты, И.О.Ф. (уч. степень, звание)	Дата подписания	Подпись
Теоретические основы автоматизации оцифровки геопространственных данных и их роль в кадастровой деятельности	PhD, ст. преподаватель Камза А.Т.	15.02.25	
Обзор и анализ технологий автоматической оцифровки	PhD, ст. преподаватель Камза А.Т.	17.04.25	
Разработка и реализация метода автоматической оцифровки на примере геологических карт	PhD, ст. преподаватель Камза А.Т.	14.05.25	
Нормоконтролер	PhD, ассоц. профессор Кенесбаева А.	17.05.25	

Научный руководитель



Камза А.Т.

Задание принял к исполнению обучающийся



Мирсидиков Ш.Х.

Дата

«17» 02 2025г

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН**

**Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И.Сатпаева»**

Горно-металлургический институт имени О.А. Байконурова

Кафедра «Маркшейдерского дела и геодезии»

Мирсидиков Шахнур Хожакбарович

**На тему: «Автоматизация процессов оцифровки и обработки геопространственных данных с
использованием современных ГИС-технологий»**

ДИПЛОМНАЯ РАБОТА

6B07304 – Геопространственная цифровая инженерия

Алматы 2025 г

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН

Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И. Сатпаева»

Горно-металлургический институт имени О.А. Байконурова

Кафедра «Маркшейдерского дела и геодезии»

ДОПУЩЕН К ЗАЩИТЕ

Заведующий кафедрой _____

Маркшейдерского дела и геодезии

к.т.н., ассоц. профессор

_____ Мейрамбек Г.

«__» _____ 2025 г.

ДИПЛОМНАЯ РАБОТА

На тему: «Автоматизация процессов оцифровки и обработки геопространственных данных с
использованием современных ГИС-технологий»

6B07304 – Геопространственная цифровая инженерия

Выполнил

Мирсидиков Ш.Х.

Рецензент

МОК

к.т.н., ассоц. профессор

_____ Омиржанова Ж.Т.

«__» _____ 2025 г.

Научный руководитель

КазННТУ им.К.И.Сатпаева

PhD, ст. преподаватель

_____ Камза А.Т.

«__» _____ 2025 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН
Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К.И.Сатпаева»

Горно-металлургический институт имени О.А. Байконурова

Кафедра «Маркшейдерского дела и геодезии»

6B07304 – Геопространственная цифровая инженерия

УТВЕРЖДАЮ

Заведующий кафедрой _____

Маркшейдерского дела и геодезии

к.т.н., ассоц. профессор

_____ Мейрамбек Г.

« » _____ 20__ г.

ЗАДАНИЕ

на выполнение дипломной работы

Обучающемуся Мирсидиков Шахнур Хожакбарович

Тема: Автоматизация процессов оцифровки и обработки геопространственных данных с использованием современных ГИС-технологий

Утверждена приказом Проректора по академическим вопросам №26-П/Ө от 29.01.2025г.

Срок сдачи законченной работы «05» июня 2025г.

Исходные данные к дипломной работе: *Материал, накопленный во время производственной практики, данные лекций, сканы геологических карт*

Краткое содержание дипломной работы:

- а) анализ теоретических основ цифровизации геопространственных данных и роли ГИС-технологий;*
- б) обзор и сравнение инструментов автоматической векторизации;*
- в) обоснование применения оцифрованных карт для нужд кадастра и инженерных изысканий;*
- г) вопросы безопасности жизнедеятельности и охраны труда; расчет экономической эффективности внедрения автоматизации.*

Перечень графического материала: *представлены 20 слайдов презентации работы*

Рекомендуемая основная литература:

1. Жеңісова Н. Е., Тұрымтаев Ж. Б., Қазықанова А. Қ., Токтаров Ж. А. Технология создания цифровой картографической основы для информационно-оценочных карт территории Республики Казахстан: диссертация / КазНУ им. аль-Фараби, 2023. – 150 с.

2. Скворцова Р.Н., Согорова И.Н. Геодезия, кадастр с основами геоинформатики: учебник для вузов / Р.Н. Скворцова, И.Н. Согорова; под ред. Р.И. Золотовой. — Алматы: Казахский техн. ун-т, 2021. — 320 с

3. Зулпыхаров К. Б. Геоинформационные системы: назначение, функции и методы векторного преобразования данных: учебное пособие / КазНУ им. аль-Фараби, 2024—200 с.

ГРАФИК
подготовки дипломной работы

Наименование разделов, перечень разрабатываемых вопросов	Сроки представления научному руководителю	Примечание
Теоретические основы автоматизации оцифровки геопространственных данных и их роль в кадастровой деятельности	15.02.2025г.	
Обзор и анализ технологий автоматической оцифровки	17.04.2025г.	
Разработка и реализация метода автоматической оцифровки на примере геологических карт	14.05.2025г.	

Подписи

консультантов и норм контролера на законченную дипломную работу с указанием относящихся к ним разделов работы

Наименования разделов	Консультанты, И.О.Ф. (уч. степень, звание)	Дата подписания	Подпись
Теоретические основы автоматизации оцифровки геопространственных данных и их роль в кадастровой деятельности	PhD, ст. преподаватель Камза А.Т.		
Обзор и анализ технологий автоматической оцифровки	PhD, ст. преподаватель Камза А.Т.		
Разработка и реализация метода автоматической оцифровки на примере геологических карт	PhD, ст. преподаватель Камза А.Т.		
9Нормоконтролер	PhD, ассоц. профессор Кенесбаева А.		

Научный руководитель _____ Камза А.Т.

Задание принял к исполнению обучающийся _____ Мирсидиков Ш.Х.

Дата «__» _____ 20__г

АНДАТПА

Бұл бітіру жұмысы геологиялық карталарды автоматты және жартылай автоматты цифрландыруға арналған бағдарламалық құралдарды салыстырмалы түрде талдауға арналған. Зерттеу нысаны ретінде үш заманауи шешім қарастырылды: ArcScan (ArcGIS Pro құрамында), Easy Tracer (QGIS плагині) және машиналық оқыту технологияларын қолданатын Bunting Labs компаниясының AI Vectorizer құралы. Зерттеудің негізгі мақсаты — дәлдігі, өнімділігі және қателерге төзімділігі тұрғысынан ең тиімді әдісті анықтау. Жұмыстың нәтижелері ГАЖ, сандық картография және геологиялық барлау саласындағы мамандар үшін практикалық маңызға ие.

АННОТАЦИЯ

Работа посвящена сравнительному анализу программных инструментов автоматической и полуавтоматической оцифровки геологических карт. В качестве объектов исследования рассмотрены три современных решения: ArcScan (в составе ArcGIS Pro), Easy Tracer (плагин для QGIS) и AI Vectorizer от Bunting Labs, основанный на технологиях машинного обучения. Целью исследования стало выявление наиболее эффективного метода с точки зрения точности, производительности и устойчивости к ошибкам. Результаты имеют прикладное значение для специалистов в области ГИС, картографии и геологоразведки.

ANNOTATION

This thesis presents a comparative analysis of software tools for automatic and semi-automatic digitization of geological maps. The study evaluates three modern solutions: ArcScan (part of ArcGIS Pro), Easy Tracer (a QGIS plugin), and AI Vectorizer by Bunting Labs, which utilizes machine learning technologies. The main objective is to identify the most efficient method in terms of accuracy, performance, and error resistance. The results offer practical value for professionals in GIS, digital cartography, and geological exploration.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	7
1 Теоретические основы автоматизации оцифровки геопространственных данных и их роль в кадастровой деятельности	8
1.1 Определение ГИС.....	8
1.2 Значимость актуальных и точных геопространственных данных для кадастра	9
1.3 Влияние автоматизации на кадастровые процессы и данные	10
2 Теоретические основы автоматизированной оцифровки геопространственных данных	13
2.1 Характеристика геопространственных данных и их роль в ГИС	13
2.2 Геологические карты как источник данных для кадастра и планирования ..	16
2.3 Методы оцифровки: ручные и автоматизированные подходы	18
2.4 Краткий обзор современных ГИС-технологий	19
2.5 Проблемы и задачи автоматизации оцифровки геологических карт	21
3 Обзор и анализ технологий автоматической оцифровки	24
3.1 Критерии выбора инструментов для автоматизации.....	24
3.2 Технология ArcScan в ArcGIS : особенности, настройки и возможности. ...	24
3.3 EasyTracer: обзор инструмента для оцифровки	30
3.4 Bunting Labs: ИИ-подходы к распознаванию картографических данных	33
3.5 Методика сравнения эффективности инструментов	36
4 Разработка и реализация метода автоматической оцифровки на примере геологических карт.....	38
4.1 Выбор геологических карт и подготовка растровых данных	38
4.2 Применение ArcScan: настройка, генерация векторов.....	39
4.4 Применение EasyTracer: загрузка, обработка, экспорт	43
4.5 Использование Bunting Labs: особенности ИИ-анализа	44
4.6 Сравнение подходов по параметрам: точность, скорость, удобство.....	46
5 Сравнительный анализ и оценка эффективности	48
5.1 Методика оценки качества оцифровки геологических карт	48
5.2 Визуальное и метрическое сравнение результатов.....	52
5.3 Сравнение полученных векторных данных с эталонными.....	55
5.4 Анализ производительности и автоматизируемости	56
ЗАКЛЮЧЕНИЕ	58
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ	59

ВВЕДЕНИЕ

Сегодня трудно переоценить влияние геоинформационных систем (ГИС) на методы обработки и анализа пространственных данных. Особенно это заметно в геологии, где точность и наглядность картографических материалов играют ключевую роль. Геологические карты, традиционно создаваемые вручную, ещё недавно хранились в бумажном формате. Однако с цифровизацией научной среды и производственных процессов остро встал вопрос: как оперативно и корректно перевести эти материалы в векторную форму, пригодную для анализа и дальнейшего моделирования в ГИС-средах?

До недавнего времени такой перевод — оцифровка — в основном осуществлялся вручную. Это отнимало массу времени и сил, а человеческий фактор нередко вносил неточности. Сейчас ситуация меняется: появляются инструменты, способные взять на себя часть этой рутинной работы. Среди них — как решения на основе классических алгоритмов, так и более продвинутые, опирающиеся на искусственный интеллект и машинное обучение. В этой работе основное внимание уделяется трем таким инструментам: ArcScan (модуль ArcGIS Pro), Easy Tracer и AI Vectorizer Bunting Labs, использующий ИИ-подход.

Актуальность темы обусловлена необходимостью критического анализа этих решений: насколько они точны, где дают сбои, в каких случаях экономят время, а в каких — только создают видимость автоматизации. В условиях стремительного роста интереса к цифровой обработке геологических данных, особенно в задачах картирования, ресурсного планирования и инженерной геологии, подобное исследование приобретает не только методическую, но и прикладную ценность.

Цель исследования — сопоставить три программных средства автоматической и полуавтоматической оцифровки геологических карт и определить, какое из них даёт наилучший результат с точки зрения точности, надёжности и производительности. Объект исследования — процесс оцифровки геологических карт в цифровую (векторную) форму. Предмет — методы и алгоритмы, лежащие в основе используемых ГИС-инструментов, а также практические особенности их применения на реальных данных.

Работа опирается на эксперименты с реальным картографическим материалом, включает визуальные иллюстрации ошибок, диаграммы точности и количественные сводки, позволяющие не только оценить результат, но и предложить гибкий, практико-ориентированный подход к выбору инструментов для задач геологической цифровизации.

1 Теоретические основы автоматизации оцифровки геопространственных данных и их роль в кадастровой деятельности

1.1 Определение ГИС

Отвечая на вопрос о будущем ГИС в грядущем тысячелетии, известный специалист по геоинформатике Нэнси Тоста выразила оптимизм: «Я убеждена, что в течение следующего десятилетия люди перестанут задаваться вопросом: «Что такое ГИС?» В этом кратком заявлении содержится глубокое понимание неизбежного внедрения географических информационных систем в повседневную профессиональную деятельность и информирование общественности.

До сих пор известные специалисты в области ГИС предложили множество определений, каждое из которых отражает специфические аспекты этой сложной технологии. В совокупности они обеспечивают целостное понимание фундаментальной природы ГИС. Например, Майкл Н. Демерс сформулировал это следующим образом: «Геоинформационные системы — это инструменты обработки пространственной информации, обычно связанной с определенными областями земной поверхности и используемой для анализа и управления ими. Как и сама география, термин «географическая информационная система» сложно свести к единому определению, поскольку он представляет собой сочетание различных дисциплин».

В нескольких публикациях ГИС описывается как система, которая:

- представляет собой сочетание пяти основных элементов: аппаратного обеспечения, программного обеспечения, пространственных данных, регламентирующих процедур и пользователей;
- предлагает полный цикл операций с географической информацией, включая ее ввод, интеграцию, хранение, обработку, анализ, моделирование и визуализацию.

Геоинформационные системы (ГИС) характеризуются рядом оригинальных определений, подчеркивающих их многогранность:

- ГИС открывают новые перспективы для анализа и понимания окружающего мира;
- ГИС служат инструментом, позволяющим рассматривать мир через призму взаимосвязей между его элементами;
- ГИС представляют собой мощное средство для интеграции и структурирования разнообразной информации.

1.2 Значимость актуальных и точных геопространственных данных для кадастра

Одним из базовых принципов ведения кадастра является точность и актуальность пространственных данных. Кадастровая карта, отображающая границы участков, должна соответствовать реальности с высокой геодезической точностью. Аналогично, любые тематические данные (в том числе геологические слои), включённые в кадастровую систему, должны быть привязаны к той же системе координат и регулярно обновляться. Международные стандарты и исследования указывают, что все координаты и планы, используемые в кадастровых работах, должны базироваться на единой геодезической основе, с тем чтобы различные слои данных корректно накладывались друг на друга. Поэтому при цифровой интеграции геологических карт необходимо выполнить их геопривязку к государственной системе координат, используемой для земельного кадастра. Только в этом случае, накладывая, например, слой геологических разломов на карту земельных участков, специалист сможет однозначно определить, какой именно участок расположен в зоне возможной сейсмической активности.

Актуальность данных подразумевает способность системы своевременно отражать изменения. Традиционные кадастры, основанные на бумажных носителях (картах, чертежах, журналах), нередко отставали от реального положения дел: обновление сведений было трудоёмким и запаздывало во времени. Геологическая информация – не статична; по мере новых исследований или событий (например, обнаружение нового разлома, изменение уровня грунтовых вод, техногенные изменения ландшафта) данные устаревают. В эпоху цифровых технологий информационно-насыщенная среда предоставляет огромные возможности для оперативного обновления данных в режиме реального времени. Современные кадастровые системы, интегрированные с ГИС, позволяют автоматически вносить изменения и сразу отображать их на картах, обеспечивая пользователей (инженеров, землепользователей, инвесторов) актуальной картиной геопространственной информации. Например, если в результате инженерно-геологических изысканий получены новые сведения о грунтах на участке, эти данные могут быть добавлены в геологический слой кадастра, и уже при подготовке технического плана или оценке участка будут учтены обновленные характеристики.

Помимо своевременного обновления, автоматизация обеспечивает и точность передачи данных. Ручное перерисовывание карты чревато человеческими ошибками – неточно проведённой линией границы, пропущенным условным знаком, неправильной подписью породы. Автоматическое же считывание и векторизация карт, особенно с применением интеллектуальных алгоритмов, позволяет значительно снизить долю субъективных ошибок. Использование высокоточных входных данных – таких как аэрофотоснимки, снимки дистанционного зондирования, данные лидар-съёмки – совместно с их обработкой в ГИС даёт возможность создавать высоко-

точные трёхмерные модели и карты. Например, совмещение результатов мобильного лидарного сканирования, георадарных исследований, аэросъёмки и сейсмических данных позволяет получить детализированную модель земельного участка и его подземных характеристик. Такие модели, интегрированные в кадастровую базу, повышают эффективность управления ресурсами и точность оценки недвижимости, одновременно снижая экологические и инженерные риски. Иначе говоря, актуальные и корректные геопространственные данные служат основой для обоснованных управленческих и проектных решений в кадастровой сфере.

1.3 Влияние автоматизации на кадастровые процессы и данные

Внедрение автоматизированных методов оцифровки геологических карт существенно влияет на эффективность и качество кадастровой деятельности. Рассмотрим основные преимущества, которые обеспечивает такая автоматизация:

Повышение точности и достоверности кадастровых данных. Автоматическая оцифровка при правильной настройке и контроле способна более точно перенести информацию с исходной карты, чем ручная отрисовка. Во-первых, устраняется субъективный фактор: алгоритм строго следует заданным параметрам (например, порогу цвета), не проводя «на глаз» линии и границы. Во-вторых, современные методы (особенно с элементами ИИ) могут распознавать мельчайшие детали, которые оператор-человек мог бы пропустить при быстром обводе. Кроме того, интеграция геологических данных в цифровой вид позволяет привязать их к точным координатам участка. Например, если ранее в кадастровом деле геологическая характеристика участка указывалась текстово (словесное описание изысканий), то теперь она может быть подкреплена цифровым геологическим слоем, отображаемым поверх границ данного участка. ArcGIS-системы, такие как ArcInfo, изначально разрабатывались для создания и ведения разнообразных кадастров – земельных, лесных, геологических и др., что свидетельствует о возможности точного учета пространственных данных разных типов в единой системе. Таким образом, автоматизация оцифровки обеспечивает интеграцию геологических карт как еще одного измерения кадастра, повышая общую точность и полноту данных об объектах.

Ускорение подготовки технической документации. Кадастровая деятельность включает значительный объем технической документации – будь то технические планы участков, отчеты по инженерно-геологическим изысканиям, акты выбора земельных участков под строительство или горнотехнические обоснования. Наличие оцифрованных геологических слоёв позволяет существенно сократить время на подготовку таких документов. Например, раньше для включения геологической информации в отчет инженеру-геологу приходилось вручную чертить фрагмент карты разреза, наносить его на план участка и т.п. Теперь достаточно выбрать нужные слои в ГИС (геология,

гидрогеология, контуры участка, строения) и сгенерировать карту или схему автоматически. При использовании шаблонов и скриптов можно автоматизировать и генерацию текстовых описаний: система будет подтягивать атрибуты из базы (например, тип грунта, категория пород) для формирования характеристик участка. Кроме того, электронные кадастровые системы поддерживают обмен данными и формирование отчетности по заданным стандартам. Так, современные решения позволяют конвертировать данные в требуемые форматы (XML, PDF с картами и т.д.) буквально в несколько кликов. В частности, опыт внедрения интегрированных кадастровых систем показывает, что одним из этапов их реализации является конверсия старых данных и унификация форматов для новой системы. В ходе этого этапа все имеющиеся сведения (в том числе геологические карты в бумаге или разрозненных файлах) собираются и интегрируются в единую базу. После такой интеграции формирование любого сводного документа (например, выписки о земельном участке с указанием геологических ограничений) становится значительно быстрее и менее трудозатратным.

Актуализация информации о геологических особенностях участков. Автоматизация связана не только с разовым переводом карт в цифру, но и с построением процессов постоянного обновления данных. Цифровой геологический слой кадастра может регулярно пополняться новыми данными: результатами свежих изысканий, данными мониторинга (например, движения грунтов), новыми картами, поступающими от геологических служб. При налаженной инфраструктуре геоданных любые изменения могут отражаться оперативно. Современные кадастровые платформы поддерживают веб-доступ и обновление данных в режиме онлайн. Это значит, что разные ведомства и специалисты (например, служба недропользования, геологическая служба, орган, ведущий земельный кадастр) могут совместно пользоваться единой системой, где обновления сразу видны всем. Автоматизация облегчает включение новых данных: если поступила цифровая карта в формате GIS, ее можно сразу загрузить; если появились новые полевые данные, они через сервисы ГИС оперативно наносятся на карту. В результате кадастр всегда содержит актуальные сведения о геологических условиях. Это важно для целого ряда практических задач: от своевременного информирования землепользователей об изменениях (например, появлении подтопления или просадки на их территории) до стратегического планирования (учёт новых разведанных месторождений при выдаче лицензий на недропользование). Например, в одном из современных проектов кадастровой системы реализована функция отображения геологического содержимого для каждого земельного участка. Пользователь, выбрав на карте интересующий участок, может сразу увидеть, какие геологические слои с ним связаны (тип грунта, категория сейсмичности, глубина залегания скального основания и т.д.), поскольку система автоматически подтягивает эти сведения из базы. Такая оперативность была бы невозможна без предварительной автоматизированной оцифровки всех релевантных геокарт и наладки их связи с единицей кадастра (участком).

Повышение эффективности межведомственного взаимодействия. Когда геологические карты оцифрованы и включены в кадастровую информационную систему, разные организации получают доступ к единому источнику данных. Земельный кадастр, горный кадастр, экологические и градостроительные службы – все работают с одной геоплатформой, где информация согласована по слоям и координатам. Это устраняет дублирование работ и рассогласованность данных. Автоматизация здесь играет ключевую роль: благодаря ей происходит унификация данных разных отраслей. Например, геологическая служба регулярно обновляет цифровой слой тектонических разломов, и эти обновления автоматически отображаются в системах городского планирования и земельного учёта. В результате при планировании нового строительства сразу учитываются новейшие данные о потенциально опасных зонах, без отдельного запроса и оцифровки бумажных карт.

Наконец, следует упомянуть и общую экономическую эффективность автоматизации. Первоначальные затраты на сканирование архивных карт, внедрение ГИС-приложений и обучение персонала окупаются за счёт ускорения кадастровых процедур и повышения качества данных. Повышение доверия к кадастру (благодаря его точности и полноте) косвенно способствует инвестиционной привлекательности территории и снижению рисков. Так, в горнодобывающем секторе наличие эффективной электронной кадастровой системы с интегрированными геоданными увеличивает приток инвесторов, а в градостроительстве – ускоряет согласования проектов за счёт прозрачности информации о условиях участка.

Подводя итог, автоматизация процессов оцифровки и обработки геологических карт является неотъемлемой частью развития современного кадастра. Современные ГИС-технологии обеспечивают платформу для слияния разнородных данных – правовых, топографических, геологических – в единое информационное поле. Автоматизированная векторизация и нейросетевые методы дают возможность быстро и с высоким качеством перевести накопленные геологические знания в цифровую форму. Включение же этих данных в кадастровую деятельность ведёт к созданию многофункциональной, точной и актуальной кадастровой системы, способной удовлетворять потребности как государства, так и бизнеса в сфере управления территориями и недрами. Таким образом, связь между автоматизацией геологических карт и кадастром проявляется двояко: с одной стороны, кадастр получает достоверную геологическую основу для решений, с другой – именно автоматизация делает возможным эффективное пополнение кадастра этой основой. Это соответствует стратегическому курсу на цифровую трансформацию земельно-ресурсной сферы и обеспечивает новый уровень качества кадастровых работ.

2 Теоретические основы автоматизированной оцифровки геопространственных данных

2.1 Характеристика геопространственных данных и их роль в ГИС

Геопространственные данные описывают «где» по отношению к «что». В отличие от стандартных данных, которые часто включают в себя списки, подсчеты или классификации, геопространственные данные включают информацию о географическом местоположении. Каждая точка данных помечена координатами, что позволяет наносить ее на карту и анализировать в контексте реальных местоположений. Это делает геопространственные данные уникальными, поскольку речь идет не только об атрибутах или переменных, но и об их расположении на поверхности Земли. Появление географических информационных систем (ГИС) в последние несколько десятилетий расширило сферу пространственных данных, что привело к созданию и использованию подробных карт и передовому пространственному анализу.

Вектор представляет собой отрезок прямой с заданным направлением от начальной точки к конечной. Его определяют начальная и конечная точки, длина и ориентация в пространстве. Векторы делятся на свободные, которые можно размещать произвольно, и несвободные, привязанные к определённому положению. Над векторами проводят операции: сложение, вычитание, а также умножение — скалярное, векторное или смешанное. Эти математические действия составляют основу для создания векторных моделей географических объектов.

В геоинформационных системах (ГИС) географические объекты отображаются на плоскости с использованием базовых графических примитивов:

- точек, представляющих собой нулевые векторы;
- линий, сформированных последовательностью двумерных векторов;
- областей, ограниченных полигонами — замкнутыми цепочками двумерных векторов. Такие объекты, представленные областями, называются полигональными.

Каждое векторное представление в цифровом формате описывается с помощью пар координат X, Y :

- географический объект, представленный точкой, задается одной парой координат (X, Y) .
- линейный объект описывается набором пар координат $(X_1, Y_1; X_2, Y_2; X_3, Y_3; \dots)$, которые формируют отрезки полилинии.
- полигональный объект представлен последовательностью координат $(X_1, Y_1; X_2, Y_2; X_3, Y_3; \dots; X_1, Y_1)$, где совпадение начальной и конечной пары указывает на замкнутость фигуры.

В геоинформационных системах (ГИС) для построения векторных моделей применяются координаты, соответствующие реальному расположению объектов на Земле. Это могут быть географические координаты (широта и

долгота) на сфероиде или плоские Декартовы координаты (абсцисса и ордината), определенные в выбранной картографической проекции.

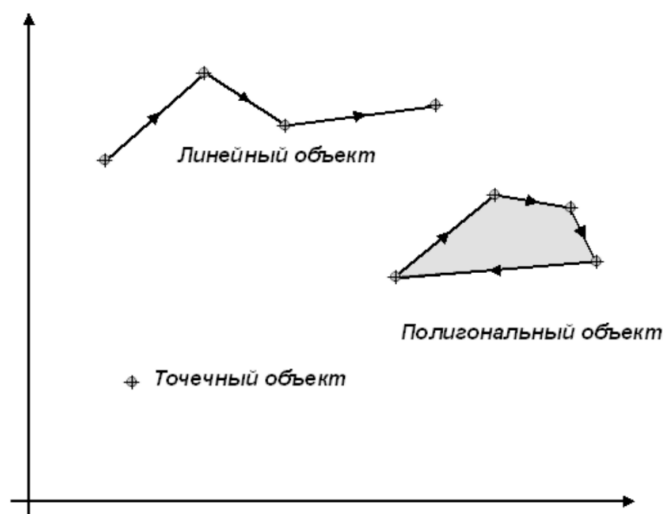


Рисунок 1 - Модели векторного представления одиночных географических объектов

Для хранения растровых данных о пространственных объектах используются различные форматы. Один из них — формат ESRI Grid, представляющий собой растровое изображение, сформированное набором пикселей.

Земная поверхность представляет собой непрерывное пространство. Для его цифрового отображения применяются модели, основанные на дискретных элементах.

Векторная модель разбивает реальное пространство на отдельные географические объекты, которые представлены в виде пространственных элементов: точек, линий или областей.

Растровая модель, в свою очередь, разделяет непрерывное пространство на сетку из смежных ячеек, являющихся основными компонентами этой модели. Ячейка — это минимальная единица информации, обладающая единственным значением или свойством, что делает её однородной. Растровая модель предоставляет данные о том, что находится в конкретной точке, тогда как векторная модель указывает на местоположение объекта. Все ячейки в растровой модели имеют одинаковый размер. В геоинформационных системах (ГИС) выделяют два типа ячеек:

- пиксель — наименьший неделимый элемент изображения.
- ячейка — элемент заданной формы (квадрат, прямоугольник, треугольник, шестиугольник или трапеция), при этом квадрат и прямоугольник наиболее распространены.

Растровые модели используются для представления непрерывных природных явлений. Растр в ГИС — это упорядоченная сетка ячеек, привязанная к реальному пространству Земли.

Наименование формата	Описание
DXF, DWG, DGN	Форматы данных систем автоматизированного проектирования (САПР)
DX90	Формат цифровых навигационных карт
DLG	Формат данных геологической съемки США
DWF	Формат передачи графических данных по Интернету
FIM	Формат данных Роскартографии, предназначен для обмена данными
GEN	Обменный формат ГИС-пакета ARC/INFO
SHP	Формат данных ГИС-пакета ArcView (шейп-файл), описывается несколькими файлами записей с определенными расширениями: .shp — позиционные данные; .shx — индекс формы пространственных данных; .dbf — атрибутивные данные и др.
TAB	Формат ГИС-пакета MapInfo; описывается файлами с определенными расширениями: .tab — текстовое описание структуры данных таблиц; .dat — табличные данные; .map — графические объекты; .ID — список указателей (индекс) на графические объекты
MIF/MID	Обменный формат ГИС-пакета MapInfo
HPGL	Формат вывода на принтер или графопостроитель
DMF	Формат ГИС-пакета Digital

Рисунок 2 - Векторные графические форматы данных

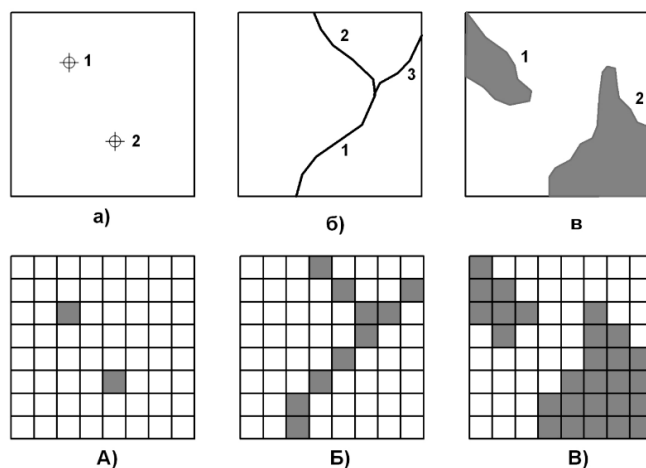


Рисунок 3 - Векторные модели (варианты а, б, в) и растровые модели (варианты А, Б, В)

Таким образом, растровая модель в ГИС представляет собой упорядоченную прямоугольную (или другую) сетку ячеек, каждая из которых содержит пространственно-привязанное значение, соответствующее определённому свойству территории. Преимуществами растровой модели являются простота структуры, высокая скорость обработки больших объёмов данных и удобство интеграции с данными спутникового мониторинга. В то же

время к её ограничениям можно отнести ограниченную точность и значительный объём памяти, требуемый для хранения высокодетализированных слоёв.

2.2 Геологические карты как источник данных для кадастра и планирования

Геологические карты содержат сведения о строении земной коры, породах, минералах, грунтах и геологических процессах, происходящих на определённой территории. Эти сведения необходимы для ведения различных видов кадастра. В земельном кадастре данные о геологическом строении участка могут влиять на его оценку, использование и ограничение в пользовании. Например, наличие плодородных почв, скальных грунтов или сейсмических разломов – всё это геологическая информация, прямо отражающая потенциальную ценность и риски при эксплуатации земли. По современным представлениям, полноценная кадастровая система должна включать инфраструктурную и геологическую информацию о каждом земельном участке. На практике же традиционные земельные кадастры долгое время ограничивались сведениями о границах и правах собственности, не включая данные о недрах и физических характеристиках земли. Это создавало пробелы в информации: без данных о геологических особенностях участок невозможно полноценно оценить ни в стоимостном, ни в градостроительном аспекте.

Горный (минеральный) кадастр напрямую зависит от геологических карт, поскольку именно на их основе определяется местоположение полезных ископаемых, границы лицензионных участков недропользования и условия добычи. Надёжные и доступные геологические данные рассматриваются как ключевой фактор привлечения инвестиций в разведку и разработку месторождений. Исследования показывают, что открытость и публичная доступность геологических карт и данных существенно повышают интерес недропользователей к региону. В странах с богатой минеральной базой внедряются электронные системы горного кадастра, интегрирующие геологические карты и сведения о недрах. Так, по опыту ряда государств, консолидация геологических данных в рамках электронного минералогосырьевого кадастра при поддержке международных программ (например, проектов Всемирного банка) способствует притоку инвестиций и повышению прозрачности управления недрами. Электронные кадастровые системы позволили потенциальным инвесторам и геологам просматривать местоположение потенциальных месторождений, геологические разрезы и границы лицензий в удобной онлайн-форме. Таким образом, для эффективного ведения горного кадастра критически важно иметь актуальные оцифрованные геологические карты, что обеспечивает информационную поддержку принятия решений в недропользовании.

Инженерно-геологические изыскания и территориальное планирование также опираются на сведения геологических карт. При проектировании зданий и

сооружений требуются данные о грунтах, глубине залегания скальных пород, уровне грунтовых вод, сейсмичности района и иных характеристиках, отражённых на геологических картах. Включение этих данных в кадастровую информацию позволяет ещё на этапе планирования выявлять ограничения: например, участки с просадочными грунтами или зонами потенциального затопления. Территориальное планирование (генеральные планы, зонирование) учитывает геологические факторы при выделении зон под застройку, сельское хозяйство или охрану природы. Наличие цифровых геологических слоёв в единой геоинформационной системе даёт возможность наложения их на кадастровые карты и топографические планы для комплексного анализа. Крупномасштабные тематические карты – геологические, почвенные, лесные – традиционно создаются государственными картографическими службами и служат базой для многих видов планов. В частности, государственные геологические карты выпускаются в стандартизованных масштабах (например, 1:50 000, 1:200 000) с унифицированной легендой и условными обозначениями. Эти официальные данные затем используются при ведении государственного земельного и горного кадастра специальными организациями. Оцифровка таких карт и включение их в цифровые базы позволяет территориальному планированию опираться на актуальные данные о недрах на всей площади планируемых территорий.

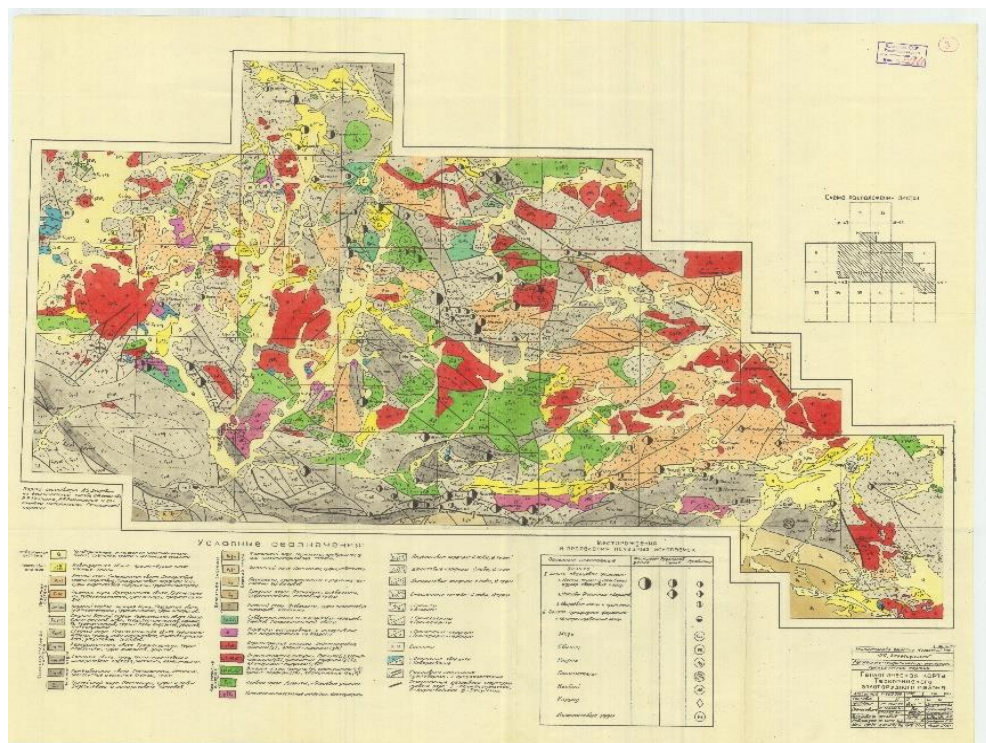


Рисунок 4 - Пример геологической карты

Таким образом, геологические карты выполняют роль важнейшего информационного слоя для системы кадастра и смежных сфер. Они обогащают кадастровые данные сведениями о природных особенностях участка, его

потенциале и ограничениях. Однако реализовать эту интеграцию на практике возможно только при условии, что геологические данные приведены в цифровую форму, сопоставимую с прочими кадастровыми сведениями (границами, координатами). Здесь и проявляется ключевая связь автоматизации оцифровки геологических карт с кадастровой деятельностью: без автоматизированных методов перевода карт в цифровой вид оперативное включение геоинформации в кадастр затруднено.

2.3 Методы оцифровки: ручные и автоматизированные подходы

Цифровизация представляет собой процесс преобразования объектов с бумажных карт в цифровую форму с использованием планшетного дигитайзера, подключённого к компьютеру. Оператор обводит элементы карты с помощью устройства, напоминающего компьютерную мышь, а их координаты (x, y) фиксируются и сохраняются как пространственные данные в виде точек, линий или полигонов. Этот процесс позволяет создавать векторные базы данных, состоящие из точечных, линейных и полигональных объектов. Ниже рассмотрены основные способы получения векторных данных, их преимущества и ограничения.

Существует четыре основных метода оцифровки: Ручная оцифровка – выполняется с помощью оцифрованного планшета. При этом методе оператор вручную обводит все линии со своей бумажной карты (например, Toposheet) с помощью указателя. Это требует много времени, но и высокой точности. Heads-up Digitizing – похож на процесс ручной оцифровки, в том смысле, что линии нужно обводить вручную, но он работает прямо на экране компьютера с использованием отсканированного растрового изображения. Уровень точности метода оцифровки выше, чем при использовании планшета для оцифровки, так как растровые изображения сканируются с высоким разрешением. Несмотря на то, что линии по-прежнему прослеживаются вручную, уровень точности выше, чем при использовании планшета для оцифровки, поскольку растровые изображения сканируются с высоким разрешением, обычно от 200 до 1600 DPI. С помощью инструментов отображения, таких как увеличение и уменьшение масштаба, оператор может фактически работать с разрешением растровых данных, следовательно, оцифровывать их с более высоким уровнем точности. Тем не менее, уровень точности все еще не гарантирован, потому что он сильно зависит от оператора и того, как он оцифровывает. Этот метод также занимает много времени и занимает примерно столько же времени, сколько и метод ручной оцифровки.

Метод интерактивного отслеживания автоматизирует процесс трассировки отдельных линий путем трассировки по одной линии за раз под руководством оператора. Это является значительным улучшением по сравнению с ручной оцифровкой с точки зрения точности и скорости оцифровки, особенно когда полностью автоматическое преобразование раstra в вектор не может быть

применено в таких случаях, как низкое качество изображения и сложные слои. Основным преимуществом использования интерактивного отслеживания является гибкость выборочного отслеживания линий и лучший контроль со стороны оператора.

Метод интерактивной трассировки – это продвинутая техника оцифровки Heads-up. В этом процессе трассировка по одной линии за раз становится более быстрой и точной.

Автоматическая оцифровка – это процесс преобразования растровых в векторные данные, который является автоматизированным методом, с использованием методов обработки изображений и распознавания образов. Он оцифровывает признаки с помощью алгоритма. Этот метод очень эффективен для работы с большими наборами данных при низких затратах времени.

2.4 Краткий обзор современных ГИС-технологий

ArcGIS представляет собой мощную и универсальную геопространственную платформу, разработанную компанией Esri, и широко признанную лидером среди решений в области ГИС. Этот инструмент обеспечивает глубокую интеграцию различных типов данных с географической составляющей, позволяя пользователям эффективно анализировать, управлять, визуализировать и обмениваться пространственной информацией.

Платформа помогает организациям принимать более обоснованные решения за счёт пространственного анализа и представления информации в контексте местоположения. ArcGIS предлагает полный набор функций и инструментов, объединённых в единую систему, способную выявлять закономерности, решать прикладные задачи и повышать эффективность бизнес-процессов.

Картографирование: Платформа позволяет по-новому взглянуть на данные с помощью их визуализации в двухмерном и трёхмерном формате, а также в режиме реального времени. Встроенные сервисы и современные картографические средства обеспечивают высокую точность и наглядность.

Пространственная аналитика: ArcGIS поддерживает расширенные аналитические возможности, включая моделирование и анализ пространственных данных, поступающих из различных источников. Геообогащение и применение алгоритмов ИИ позволяют существенно углубить уровень аналитики.

Работа в полевых условиях
Платформа обеспечивает мобильным командам инструменты для сбора и обработки данных в офлайн-режиме, помогает оптимизировать логистику, маршрутизацию и управление выездным персоналом.

Управление и редактирование данных
ArcGIS предлагает стабильную и надёжную среду для обновления и координации

пространственных данных, включая возможности автономной работы и совместного доступа к информации на всех уровнях организации.

Анализ изображений и дистанционное зондирование Система включает в себя мощный инструментарий для обработки и анализа спутниковых и аэрофотоснимков. Поддерживаются автоматизированные рабочие процессы с применением ИИ, что ускоряет обработку изображений и делает интерпретацию данных более точной.

QGIS — это мощная геоинформационная система (ГИС) с открытым исходным кодом, развитие которой началось в мае 2002 года как проект на платформе SourceForge. Сегодня QGIS активно используется на различных операционных системах, включая Unix, Windows и macOS. Разработка ведётся на языке C++ с применением кроссплатформенного фреймворка Qt, что обеспечивает высокую производительность и интуитивно понятный графический интерфейс.

Среди ключевых достоинств QGIS — гибкость в работе с широким спектром пространственных данных без необходимости предварительного преобразования в единый формат. Программа поддерживает как векторные, так и растровые данные в 2D и 3D-представлении, обеспечивая их совместное использование в различных проекциях.

Система работает с данными из множества источников:

- Векторные форматы, включая GeoPackage, ESRI Shapefile, MapInfo, SDTS, GML и другие, поддерживаемые через библиотеку OGR.
- Пространственные базы данных, такие как PostGIS, SpatiaLite, MS SQL Spatial и Oracle Spatial.
- Растровые изображения и форматы, обработка которых осуществляется через GDAL: GeoTIFF, ERDAS IMG, ArcInfo ASCII GRID, JPEG, PNG и пр.
- Поддержка трёхмерных моделей, включая TIN и регулярные сетки.
- Векторные тайлы и данные GRASS GIS (как растровые, так и векторные).
- Пространственные веб-сервисы по стандартам OGC — WMS, WMTS, WCS, WFS и WFS-T.
- Табличные данные в форматах ODS и XLSX.
- Также имеется поддержка временных данных, что особенно важно для динамического моделирования и анализа.

Таким образом, QGIS предлагает обширный инструментарий для специалистов, работающих с пространственными данными, и представляет собой эффективную альтернативу коммерческим ГИС-решениям.

Easy Trace — это специализированное программное обеспечение для векторизации, предназначенное для расшифровки и обработки как хорошо сохранившихся, так и изношенных карт и снимков. Его основное назначение — преобразование аналоговых источников в современные цифровые данные для использования в ГИС. Благодаря универсальности, программа подходит для подготовки данных практически под любую ГИС-платформу, а продуманный интерфейс обеспечивает высокую эффективность даже при полностью ручной оцифровке.

В основе архитектуры Easy Trace лежит концепция мозаичного растрово-векторного поля, которое практически не имеет ограничений по размеру. Отдельные растры могут достигать более 2 Гб, быть разной глубины цвета и масштабов, формируя многослойную мозаичную структуру. Количество векторных слоёв не ограничено, и каждый из них может включать до миллиона объектов, что делает возможным построение детализированных карт, охватывающих, например, целые городские территории.

Особое внимание в Easy Trace уделено удобству работы: интерфейс программы собрал в себе лучшие черты профессиональных систем ГИС и САПР. Он одновременно интуитивно понятен для оператора и предоставляет руководителям проектов инструменты для контроля качества и интеграции данных.

Программа функционирует в среде Windows (версии 95/98/2000/NT) и предназначена для полуавтоматической интерактивной векторизации как цветных, так и монохромных растров. Easy Trace особенно полезен при переводе бумажных картографических материалов в цифровой формат, обеспечивая надёжную и точную передачу графической информации в векторный вид.

Таким образом, Easy Trace служит надёжным инструментом для специалистов, работающих с пространственными данными, и значительно облегчает переход от аналоговых источников к полноценным ГИС-решениям.

2.5 Проблемы и задачи автоматизации оцифровки геологических карт

В последние годы автоматизация перевода геологических и географических данных в цифровой формат становится всё более актуальной задачей в рамках геоинформационных исследований. Геологические карты по-прежнему играют ключевую роль как источник сведений о строении земной коры, распределении горных пород и месторождениях полезных ископаемых. Их цифровизация открывает новые горизонты: от пространственного анализа до интеграции в сложные ГИС-платформы. Тем не менее, несмотря на успехи в развитии технологий машинного зрения и нейросетей, полностью автоматизировать этот процесс пока не удаётся — в силу ряда технических и методических препятствий.

Наиболее острые проблемы автоматизации:

- Сложная визуальная композиция. Геологическая карта — это не просто график или чертёж. В ней переплетаются литологические границы, цветовые пятна, штриховки, надписи и условные обозначения. Эти элементы нередко накладываются друг на друга, что серьёзно усложняет задачу автоматического распознавания.

- Качество исходных изображений. В реальности обработка часто начинается не с идеального изображения, а со сканов старых бумажных карт. Зернистость, размытие, артефакты и потёртости создают дополнительный шум, который мешает точной интерпретации данных.

- Непостоянство оформления. Между картами, созданными в разные эпохи и разными организациями, существуют заметные различия — от цветовой гаммы до структуры легенд и набора символов. Это ставит под сомнение возможность создания универсального алгоритма, одинаково эффективно справляющегося с любыми материалами.

Автоматизация процесса оцифровки геологических карт остаётся сложной задачей, несмотря на значительные достижения в области нейросетей и компьютерного зрения. Геологические карты являются ключевым источником информации о строении земной коры, распределении пород, тектонических нарушениях и месторождениях полезных ископаемых. Их перевод в цифровую форму позволяет интегрировать пространственные данные в ГИС, проводить аналитическое моделирование и создавать современные цифровые геоданные. Однако технологические и методологические ограничения по-прежнему затрудняют создание универсальных и устойчиво работающих систем автоматической векторизации. Сложности начинаются уже на этапе восприятия изображения: геологические карты отличаются сложной многослойной структурой, где графические элементы, текст и цветовые заливки нередко перекрываются и смешиваются. Распознавание таких композиций требует сложных алгоритмов сегментации и предварительной очистки изображения, что особенно критично при работе со старыми отсканированными материалами, подверженными шумам, потерям и искажениям.

Одной из существенных проблем остаётся отсутствие обширных размеченных наборов данных, необходимых для обучения нейросетевых моделей. Ограниченность доступа к картографическим материалам, особенно архивным, а также необходимость в участии геологов-экспертов при разметке делает создание обучающих датасетов трудоёмким и дорогим процессом. Это, в свою очередь, замедляет развитие универсальных решений, способных работать с картами разных стилей и эпох. Существующие инструменты, такие как ArcScan, хорошо справляются с простыми случаями, где объекты имеют чёткие геометрические формы и однородные заливки, однако теряются в сложных случаях с плавными литологическими границами, смешанными цветами и неструктурированной символикой. Попытки применить типовые алгоритмы бинарной сегментации или пороговой фильтрации часто приводят к распознаванию ошибок, дублированию контуров или потере связности объектов, что требует последующей ручной корректировки — зачастую более трудоёмкой, чем прямое векторизование вручную.

Таким образом, ключевые задачи в области автоматизации можно сгруппировать по нескольким направлениям. Во-первых, требуется разработка устойчивых систем предобработки изображений, которые смогут удалять шумы, повышать резкость, балансировать контраст и избирательно выделять элементы, значимые для последующей обработки. Во-вторых, необходимо внедрение интеллектуальных систем распознавания и сегментации, способных классифицировать участки карты по цвету, текстуре и структуре, используя как классические алгоритмы, так и современные архитектуры нейросетей, такие как

U-Net. Особое значение приобретают методы машинного обучения с частичным контролем, включая полунаблюдаемое обучение и transfer learning, которые позволяют обучать модели даже при ограниченном объёме аннотированных данных. Кроме того, важно обеспечить корректное автоматическое преобразование растровых данных в векторные с сохранением топологии — то есть соблюдением пространственной структуры объектов, их связей и последовательности, особенно при подготовке карт к геопривязке и дальнейшему анализу в ГИС-среде.

Наконец, особое внимание следует уделить пользовательским интерфейсам, позволяющим осуществлять валидацию и редактирование результатов векторизации. Такая система должна быть встроена в знакомую среду (например, QGIS или ArcGIS Pro), обеспечивать удобный просмотр, быстрое редактирование, а также возможность отслеживания качества преобразования. Для объективной оценки эффективности алгоритмов необходима формализация метрик — таких как точность (precision), полнота (recall), визуальное соответствие оригиналу и корректность топологических связей. В совокупности все эти направления формируют основу для построения надёжной, масштабируемой и действительно полезной системы автоматизации оцифровки геологических карт, соответствующей требованиям современной геоинформатики.

Таблица 1 - Проблемы автоматизации оцифровки геологических карт

Категория проблемы	Соответствующая задача автоматизации
Низкое качество исходных карт	Улучшение алгоритмов предобработки и фильтрации
Разнородность символики	Унификация стилей и обучение на выборках
Проблемы с OCR текста	Повышение точности распознавания текста (OCR)
Топологические ошибки	Реализация автокоррекции топологии
Ограничения ПО	Разработка адаптивных инструментов
Отсутствие полной автоматизации	Максимальная автоматизация с контролем качества

3 Обзор и анализ технологий автоматической оцифровки

3.1 Критерии выбора инструментов для автоматизации

Автоматизация процессов оцифровки и обработки геопространственной информации — ключевой этап в цифровизации пространственных данных, особенно в таких сферах, как кадастр, геология, картография и территориальное планирование. Успех этих процессов во многом зависит от корректного выбора программных инструментов, обеспечивающих точную, быструю и масштабируемую трансформацию растровых карт в векторный формат.

С учётом цели дипломной работы — анализа и применения современных ГИС-технологий — выбор программных средств оцифровки должен основываться на следующих критериях:

1. Точность векторизации — автоматизированный инструмент должен обеспечивать высокую точность распознавания геометрических объектов (линий, контуров, символов).
2. Гибкость настройки — наличие параметров адаптации под конкретные типы карт (толщина линий, уровень шума, цветовые каналы).
3. Скорость обработки — высокая производительность при работе с массивами данных или крупномасштабными картами.
4. Интеграция с ГИС-платформами — совместимость с QGIS, ArcGIS, MapInfo и другими средами пространственного анализа.
5. Форматы вывода — поддержка открытых форматов: *.shp, *.geojson, *.gpx, *.dxf и т.д.
6. Интерфейс и пользовательский опыт — доступный интерфейс, наличие локализации, возможность визуального контроля результатов.
7. Искусственный интеллект и машинное обучение — использование нейросетей для повышения качества оцифровки, особенно в работе со сканами и шумными изображениями.
8. Лицензирование и поддержка — open-source или коммерческая модель, наличие поддержки, документации и пользовательского сообщества.

Выбор инструмента оцифровки должен быть обусловлен задачами проекта, типом исходных данных и требованиями к точности. В рамках данной дипломной работы сравниваются возможности и ограничения нескольких решений (ArcScan в ArcGIS, Easy Tracer в QGIS и AI Vectorizer), что позволяет определить наиболее эффективные подходы для использования в кадастровой и инженерно-геологической практике.

3.2 Технология ArcScan в ArcGIS : особенности, настройки и возможности

Модуль ArcScan, входящий в состав ArcGIS, представляет собой мощный инструмент, предназначенный для преобразования растровых изображений в

векторные слои. Эта технология особенно востребована при необходимости оцифровки традиционных картографических материалов, сохранившихся в аналоговом формате.

Процесс, известный как векторизация, может быть выполнен двумя основными способами: вручную — через интерактивную трассировку растра, когда пользователь последовательно обводит пиксельные элементы изображения, или в автоматическом режиме, при котором векторные объекты создаются на основе заранее заданных параметров. Интерактивная векторизация позволяет более точно контролировать результат, но требует больше времени и усилий, тогда как автоматическая векторизация даёт быстрый результат и хорошо подходит для карт с чёткой графической структурой. ArcScan делает возможным оба подхода, предоставляя пользователю гибкость в выборе метода в зависимости от качества исходных данных и поставленных задач. Применение данной технологии особенно актуально для организаций, которым необходимо интегрировать бумажные географические материалы в цифровую среду — будь то инженерные чертежи, топографические планы или другие картографические документы.

Благодаря ArcScan, процесс трансформации старых карт из растрового в векторный формат становится значительно проще и эффективнее по сравнению с традиционными методами, такими как использование дигитайзеров. Кроме того, модуль включает инструменты для подготовки растровых изображений перед векторизацией, позволяя удалить лишние элементы, не относящиеся к нужной информации. Такая предварительная обработка, известная как очистка растра, критически важна для повышения точности результата: удаление шумов, избыточных линий или артефактов позволяет улучшить структуру данных и минимизировать количество ошибок при векторизации.

В целом, ArcScan представляет собой надёжное и функционально гибкое решение для специалистов, работающих на стыке аналоговой картографии и цифровых геоинформационных систем.

Начало работы с ArcScan: модуль ArcScan интегрируется непосредственно в среду ArcMap и представляет собой специализированный интерфейс, обеспечивающий доступ к инструментам и командам, необходимым для выполнения процесса векторизации. Как и в случае с другими функциональными модулями ArcGIS, перед началом работы необходимо активировать ArcScan через меню настроек, а также добавить соответствующую панель инструментов на карту проекта, чтобы получить возможность управлять параметрами и использовать доступные функции. Поскольку векторизация с помощью ArcScan осуществляется в контексте редактирования, пользователю потребуется запустить сеанс редактирования (Start Editing), что открывает доступ к полноценному функционалу редактирования и позволяет комбинировать возможности ArcScan с базовыми средствами редактирования объектов.

В процессе работы используются такие настройки, как параметры замыкания линий, шаблоны целевых объектов и активные векторные слои, в которые будут записываться новые данные. Важно отметить, что для выполнения

автоматической векторизации ArcScan предъявляет определённые требования к исходному растру — он должен быть представлен в виде двухцветного изображения. Это означает, что необходимо назначить два уникальных цветовых значения растровому слою, обычно соответствующие изображению и фону. Как правило, картографические документы имеют тёмный передний план (например, чёрные линии или символы) и светлый фон (белый), однако возможна и обратная цветовая схема — главное, чтобы цвета были различимыми и уникальными для распознавания. Для задания нужных цветов можно воспользоваться режимами отображения раstra «Уникальные значения» (Unique Values) или «Классификация» (Classified), доступными в свойствах слоя. Процесс преобразования растровой графики в векторные объекты осуществляется на основе заданных пользователем параметров, влияющих на геометрию итоговых объектов. Настройки векторизации можно адаптировать под конкретный тип карты, сохранить в шаблон и использовать повторно, что особенно удобно при работе с сериями однотипных изображений или при стандартизации рабочих процессов в рамках одного проекта.

Таким образом, ArcScan обеспечивает гибкий и контролируемый подход к векторизации, позволяя эффективно интегрировать аналоговые картографические материалы в цифровую ГИС-среду.

Интерактивная векторизация, также называемая трассировкой раstra, представляет собой процесс, при котором пользователь вручную управляет созданием векторных объектов на основе растрового изображения. Этот подход оказывается особенно полезным в ситуациях, когда требуется векторизовать небольшие участки карты или когда автоматическая векторизация не обеспечивает необходимого качества. По своей сути метод схож с традиционным редактированием объектов в ГИС, однако в данном случае создаваемые объекты ориентируются непосредственно на структуру растрового слоя.

Одной из ключевых функций при трассировке является возможность «замыкания» на элементы раstra. Хотя замыкание не является обязательным, его использование значительно повышает точность и контролируемость результата. Система позволяет «привязываться» к различным элементам изображения: центральным линиям, углам, пересечениям, конечным точкам и сплошным областям. Это упрощает точное воспроизведение исходных форм, особенно при векторизации сложных геологических структур или инженерных чертежей.

Основным инструментом, реализующим данный процесс, является команда Vectorization Trace. С её помощью пользователь может вручную обводить ячейки раstra, формируя линейные или полигональные векторные объекты. Достаточно указать направление трассировки курсором и щелкнуть мышью — при каждом таком действии программа создаёт объекты вдоль ячеек исходного изображения. Этот метод особенно эффективен при использовании в сочетании с замыканием, так как позволяет добиться более точного следования линии.

На результат векторизации напрямую влияют текущие параметры, заданные пользователем. Эти настройки определяют геометрию создаваемых

объектов, включая сглаживание, точность следования и тип (линия или полигон). Перед началом работы рекомендуется отрегулировать параметры векторизации в соответствии с характером обрабатываемого изображения. Кроме того, итоговые объекты можно доработать с помощью инструментов редактирования ArcGIS, включая средства топологической коррекции, трансформации и другие функции для уточнения геометрии.

Таким образом, трассировка растра является гибким инструментом, обеспечивающим высокий уровень контроля над процессом векторизации. Она особенно полезна в тех случаях, когда автоматические алгоритмы не справляются с распознаванием, а также при работе с фрагментами карт, требующих точного соответствия оригиналу.

Распознавание формы - одной из полезных возможностей, предоставляемых ArcScan, является функция распознавания формы. С её помощью пользователь может автоматически выделять и создавать векторные объекты на основе растровых элементов, имеющих определённую геометрию. Например, такие объекты, как здания или резервуары, могут быть мгновенно распознаны и векторизованы одним щелчком мыши. Это особенно удобно при работе с картами, содержащими множество повторяющихся элементов стандартной формы. Технология существенно сокращает время ручной оцифровки и минимизирует вероятность ошибок, связанных с неточной отрисовкой контуров.

Автоматическая векторизация представляет собой технологию, позволяющую преобразовать растровые данные в векторные объекты без необходимости ручного вмешательства на каждом этапе. Несмотря на автоматический характер, процесс полностью контролируется пользователем, который задаёт параметры и влияет на конечный результат. Среди ключевых факторов, определяющих качество векторизации, выделяются разрешение изображения, уровень шума и сложность графической структуры сканированного документа. От этих характеристик зависит, насколько точно и корректно будут распознаны объекты и преобразованы в векторный формат.

ArcScan предлагает два основных метода автоматической векторизации: по центральной линии и по контуру. Первый способ используется для генерации векторных объектов, расположенных вдоль середины линейных элементов растрового изображения — он особенно эффективен для оцифровки дорожной сети, границ и инженерных коммуникаций. Второй метод — контурная векторизация — формирует объекты по внешним границам элементов, что позволяет более точно воспроизвести очертания площадных структур. Выбор между этими способами зависит от характера исходной карты и требований к результату. Оба метода могут быть настроены через интерфейс ArcScan, что позволяет добиться максимального соответствия исходным данным.

Для выполнения автоматической векторизации в ArcScan требуется предварительная настройка параметров, определяющих способ формирования векторных объектов. Эти параметры, известные как стили векторизации, позволяют точно задать правила преобразования и могут быть сохранены для

повторного использования при обработке других растров с аналогичными характеристиками. Такая повторяемость особенно удобна при пакетной оцифровке серии карт, выполненных в одном стиле.

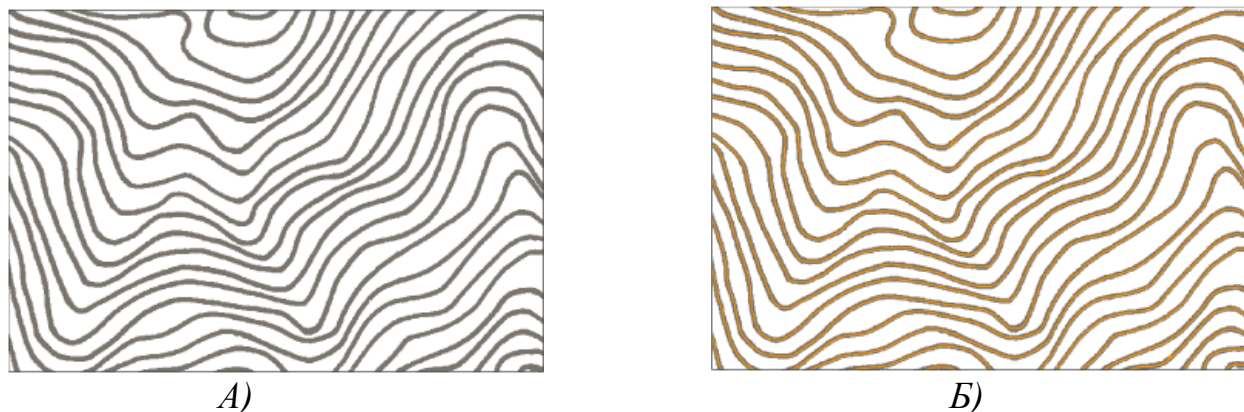


Рисунок 5 – А) Растр, Б) Вектор

Качество и успешность автоматической векторизации во многом зависят от состояния исходного растрового изображения. Иногда требуется внести изменения до начала векторизации — например, удалить шум, лишние элементы или подчеркнуть границы объектов. Этот этап, называемый предварительной обработкой растра, играет ключевую роль в достижении корректного результата. Для этих целей в ArcScan предусмотрен набор инструментов Raster Cleanup, позволяющий очистить или откорректировать отдельные участки изображения. Дополнительно можно использовать так называемые растровые выборки — они помогают изолировать конкретные ячейки, подлежащие векторизации, как в рамках всего изображения, так и независимо от него.

Одним из важнейших аспектов настройки автоматической векторизации являются параметры генерализации и сглаживания. Именно они определяют, насколько точно итоговые векторные объекты будут следовать исходной растровой структуре и насколько допустимо упрощение геометрии. Пользователь может в режиме реального времени изменять параметры и наблюдать, как это влияет на форму и точность создаваемых объектов. Это даёт возможность подобрать оптимальные значения для конкретной задачи.

После завершения векторизации как всего слоя, так и отдельных его фрагментов, результат можно доработать с использованием стандартных инструментов ArcGIS. Среди них — функции топологической коррекции, расширенного редактирования и геометрической трансформации. Это особенно актуально в случаях, когда требуется довести итоговые данные до уровня, соответствующего профессиональным стандартам пространственного анализа и картографической визуализации.

Выбор ячеек растра: ArcScan предоставляет пользователю возможность избирательно работать с растровыми ячейками, что открывает дополнительные сценарии использования при подготовке данных к векторизации. Под смежными

ячейками понимаются пиксели, имеющие непрерывные границы — они могут находиться как в прямом соприкосновении, так и соединяться по диагонали. Эта функциональность позволяет пользователю выделять фрагменты растра, с которыми предполагается дальнейшая работа: от векторизации до экспорта или удаления. Выбор конкретных областей изображения позволяет более точно управлять процессом обработки данных и исключать из анализа те фрагменты, которые не представляют интереса или могут исказить результат.

Для повышения точности и сосредоточения на значимых частях изображения применяется инструмент растровой выборки. Он позволяет изолировать необходимые области, минимизируя вмешательство посторонних элементов. Создать выборку можно двумя основными способами: щелчком мыши по взаимосвязанным пикселям или с помощью запросов, сформированных на основе логических выражений. Эти методы особенно полезны при подготовке области для последующей векторизации, поскольку позволяют заранее задать границы обработки и исключить нерелевантные участки изображения.

Очистка растров до векторизации - дополнительно ArcScan предлагает набор инструментов для редактирования растровых изображений непосредственно в процессе сеанса редактирования в ArcMap. Пользователь может рисовать, закрашивать или стирать ячейки, тем самым удаляя или добавляя элементы в изображение перед его преобразованием в векторную форму. Этот этап подготовки, известный как очистка растра, позволяет исключить пиксели, не подлежащие векторизации, и, таким образом, повысить точность и чистоту результирующего векторного слоя. При необходимости изменённое изображение можно сохранить как отдельный файл, не затрагивая оригинал, что особенно важно для обеспечения целостности исходных данных и ведения архивов. Эти инструменты позволяют гибко подстраивать данные под требования конкретной задачи, повышая эффективность последующего анализа в ГИС-среде.

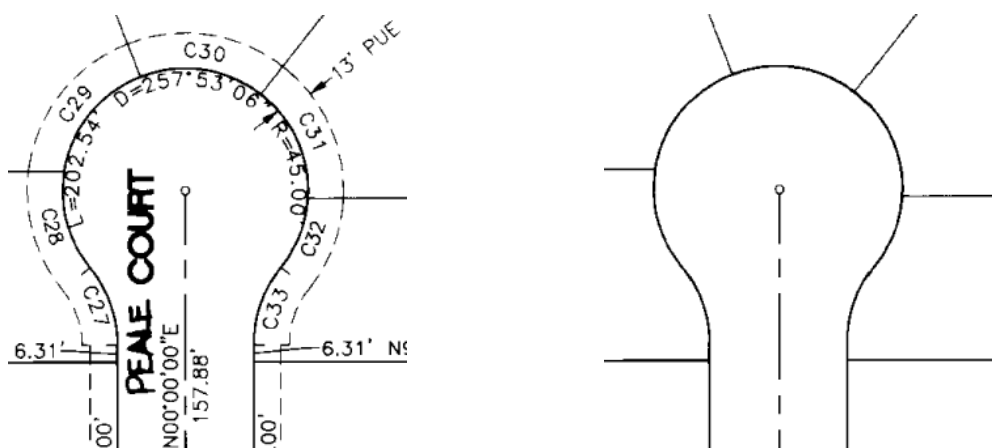


Рисунок 6 - до и после очистки растра

3.3 Анализ возможностей инструмента EasyTracer для автоматизированной оцифровки

Easy Trace представляет собой универсальное программное средство, предназначенное для решения широкого спектра задач, связанных с векторизацией растровых изображений. Его интерфейс и функциональность ориентированы как на начинающих пользователей, так и на профессионалов, работающих в области картографии, геоинформатики и САПР. Ключевыми направлениями применения Easy Trace являются оцифровка черно-белых, полноцветных и индексированных изображений. Программа позволяет выполнять векторизацию как в исходной цветовой гамме, так и с использованием встроенных средств цветоделения и замены цветов, что особенно полезно при работе с картами сложного оформления.

Среди функциональных возможностей Easy Trace можно выделить поддержку прямого сканирования из интерфейса программы с использованием сканеров различных форматов, включая настольные модели. Для подготовки изображения к векторизации предусмотрен полный набор инструментов: поворот, обрезка, сшивка фрагментов, фильтрация шумов, выравнивание и геометрическая коррекция. Встроенные средства трансформации обеспечивают коррекцию искажений с использованием кусочно-проективных, аффинных и квадратичных преобразований, что позволяет повысить точность итоговых данных и устранить деформации, возникшие при сканировании.

Одним из преимуществ Easy Trace является возможность автоматического прослеживания объектов с интерактивным контролем в процессе векторизации. Это особенно актуально при работе с изображениями низкого качества или с высокой степенью визуальной сложности. Программа поддерживает оцифровку различных типов линий (включая пунктирные, ортогональные и специальные), автоматическое оконтуривание залитых и заштрихованных участков, а также полуавтоматическую векторизацию 3D-полилиний, где каждой точке может быть присвоена своя высота (Z-координата).

Easy Trace также позволяет строить иерархическую структуру слоёв, распределяя векторные объекты по логическим уровням. При этом учитываются требования конечных ГИС и САПР-систем: реализована автоматическая генерация узлов в точках пересечения линий, замыкание полигонов, копирование совпадающих участков между слоями. Программа включает инструменты верификации топологии — проверку на пересечения, висячие вершины, незамкнутые контуры и другие ошибки, что обеспечивает чистоту и структурную целостность данных.

Дополнительно предусмотрена поддержка условных знаков и библиотек символов, упрощающих ввод и контроль корректности графической информации. Easy Trace также предоставляет инструменты для работы с базами данных DBF-формата: можно задать отдельные структуры для точечных и линейных объектов, связать их с геометрией через пользовательские идентификаторы, заполнять таблицы в процессе векторизации и автоматически

передавать Z-координаты. Это делает возможным полноценное управление как векторной, так и атрибутивной частью данных.

Результаты векторизации могут быть экспортированы в форматы популярных ГИС и САПР-платформ, включая AutoCAD, ArcCAD, Arc/Info, Intergraph, MapInfo и CREDO. Программа также поддерживает импорт векторных данных из этих систем, что обеспечивает двухстороннюю совместимость и интеграцию с существующими рабочими процессами. При необходимости рабочий проект может быть распределён между несколькими пользователями, а результаты объединены в единый векторный слой, что особенно удобно при масштабных задачах.

В обновлённой версии Easy Trace реализован ряд функциональных улучшений, направленных на повышение удобства и точности работы с растровыми и векторными данными. Одним из нововведений стала возможность прямой печати проектов непосредственно из среды Easy Trace, что упрощает подготовку отчётной документации и визуализацию результатов оцифровки.

В пакет добавлен инструмент «Кисть» для локального редактирования растра. Он позволяет закрашивать выбранные участки изображения в окрестности курсора заданным цветом, что особенно полезно при корректировке растрового слоя перед векторизацией. Эта функция повышает точность распознавания объектов и помогает избежать ошибок при автоматической трассировке.

Функционал автоматической трассировки был существенно расширен. Теперь пользователь может выполнять трассировку выбранных растровых слоёв с возможностью настройки уровня точности, типа объектов (линии или полигоны), а также выбора режима фильтрации. Добавлена функция автоматической сшивки концов линий в пределах заданного радиуса с использованием многокритериального контроля, включающего анализ различий по ширине линий, Z-координатам, атрибутивным характеристикам и принадлежности к тематическим слоям.

В систему интегрированы функции фильтрации, сглаживания и оптимизации формы линий. Эти инструменты полезны не только при автоматической векторизации, но и при последующей доработке полуавтоматически введённых объектов, обеспечивая более корректное и «чистое» отображение данных.

Инструменты полуавтоматической трассировки были полностью переработаны: усовершенствованы алгоритмы, улучшен пользовательский интерфейс и расширены возможности визуального контроля над процессом. Обновлённые методы лучше восстанавливают форму линий, эффективно обрабатывают участки с утолщением и обеспечивают более гибкие стратегии оцифровки.

Для трассировки простых геометрических фигур, таких как окружности и прямоугольники, реализован полуавтоматический режим, в котором объект можно оцифровать буквально одним щелчком мыши. Кроме того, в инструментах полуавтоматической трассировки линий теперь доступны два режима работы:

отслеживание по центру линии и оконтуривание. Отдельный инструмент для оконтуривания был удалён как избыточный — его функциональность перенесена в общий механизм.

Существенно изменилась логика настройки инструментов трассировки: теперь они настраиваются не как отдельные параметры, а как стратегии обработки конкретных типов объектов — например, зданий, рек, дорог или границ. Выбор стратегии определяет используемый инструмент автоматически, что ускоряет настройку проекта и делает интерфейс более интуитивно понятным.

Для повышения точности редактирования векторных данных в пакет внедрён новый редактор топологии. Он поддерживает операции по изменению местоположения узловых точек — при переносе одного узла автоматически смещаются все связанные с ним линии. Также доступно редактирование общих вершин, при котором система отслеживает все точки, образующие одно геометрическое соединение. В случае удаления последних вершин, например, в линиях из одной точки или полигонах с тремя вершинами, система автоматически удаляет сам объект, предотвращая появление некорректных элементов в данных.

Среди новых инструментов, реализованных в Easy Trace, стоит отметить появление функции «Пипетка», которая позволяет формировать цветовые наборы при работе с растровыми изображениями любой глубины и цветности. Это значительно облегчает процесс цветоделения, особенно в задачах, связанных с сегментацией многослойных карт. Кроме того, было снято прежнее ограничение на количество векторных объектов в одном слое — теперь один слой способен содержать до миллиона элементов, что особенно важно при оцифровке крупномасштабных и комплексных картографических данных.

Заметно улучшена производительность при открытии проектов. Реализована функция фоновой загрузки растров: пользователь может приступить к работе с векторной информацией (редактирование, просмотр, экспорт и пр.) ещё до полной подгрузки всех проектных изображений. Это особенно полезно при работе с тяжёлыми или многостраничными растровыми наборами. Также устранено ограничение по объёму обрабатываемых файлов: теперь программа способна работать с растрами размером свыше 2 Гб, вплоть до пределов, определяемых свободным местом на диске пользователя.

Добавлена полноценная поддержка форматов TIFF и JPEG, что позволяет экономить место при сохранении полноцветных изображений без потери качества. Нововведением стало и возможность редактирования растра с применением интерполяции, благодаря чему при трансформации исходного изображения значительно улучшается его визуальное качество. Введена функция отмены действий (Undo) для операций с растром, что расширяет свободу пользователя и позволяет экспериментировать без риска потери данных.

Инструмент «Линейка» был переработан и теперь применяется как к растровым, так и к векторным объектам. Расширена система навигации по проекту: реализована команда включения и выключения растров, находящихся под заданной точкой — особенно полезная функция при работе с проектами,

содержащими множество слоёв. Для произвольного перемещения по полю проекта добавлен инструмент «Рука», а также представлено новое окно «Навигатор», отображающее всю область проекта и текущее положение активного окна. Благодаря «Навигатору» стало возможным быстро переключаться между участками проекта простым указанием нужной зоны на мини-карте.

Для повышения удобства настройки программы были внедрены новые элементы управления, которые чётко указывают единицы измерения и допустимый диапазон изменяемого параметра. Это существенно упростило процесс конфигурации интерфейса и сделало программу более доступной не только для опытных пользователей, но и для начинающих специалистов.

3.4 Bunting Labs: ИИ-подходы к распознаванию картографических данных

В период стремительного развития технологий искусственного интеллекта, особенно в 2023 году, специалисты в области геоинформационных систем всё чаще стали рассматривать машинное обучение как способ автоматизировать рутинные, ресурсоёмкие процессы. Одной из таких задач, традиционно отнимающей большое количество времени, остаётся оцифровка растровых изображений. В начале 2024 года компания Bunting Labs представила плагин для QGIS — AI Vectorizer, предлагающий свежий взгляд на решение этой проблемы.

Основная идея плагина заключается в том, что он не заменяет пользователя, а дополняет его действия. Технология, которую разработчики называют автозаполнением, реализована как интерактивное продолжение линий: пользователь начинает отрисовку векторного объекта, а модель на основе визуального контекста дорисовывает оставшийся фрагмент. При этом пользователь сохраняет полный контроль над процессом и может в любой момент остановить или скорректировать результат. Такой подход делает оцифровку гибкой и управляемой, сводя к минимуму необходимость ручного труда при сохранении высокого уровня точности.

В отличие от более универсальных моделей, таких как Segment Anything, AI Vectorizer не стремится понять, что изображено на карте. Segment Anything показывает хорошие результаты при работе с чёткими спутниковыми снимками или аэросъёмкой, где можно, например, легко выделить границу водоёма по простому текстовому запросу. Однако такие модели начинают испытывать трудности при работе с историческими или плохо отсканированными картами, где смысл линий понятен только специалисту. Машинное определение значимости отдельных элементов на таких изображениях становится крайне ненадёжным.

Разработчики Bunting Labs подошли к проблеме иначе: вместо того чтобы пытаться научить модель «понимать» карту, они передали это понимание самому пользователю. В этом заключается ключевое отличие AI Vectorizer от

большинства существующих решений. Он не анализирует значения линий, а фокусируется исключительно на визуальной продолжительности начатой пользователем геометрии. Таким образом, пользователь задаёт направление и форму, а система помогает довести начатое до конца — быстро и без потери контекста.

Механизм автоматического дополнения в процессе цифровизации функционирует следующим образом: пользователь инициирует создание линии или полигона на растровом изображении, задавая начальные две точки. После этого небольшой участок карты отправляется на сервер, где алгоритм анализирует начерченный фрагмент и определяет траекторию для последующих вершин, автоматически добавляя до 50 точек. Этот подход эффективен для любых типов линий, включая прерывистые или пересекающиеся с другими элементами. В случае отклонения траектории от желаемого пути пользователь может отменить последний этап, выбрав последнюю корректную вершину, что минимизирует временные потери и позволяет продолжить работу.

Для работы с плагином "Bunting Labs AI Vectorizer" в среде QGIS необходимо сначала зарегистрироваться, указав адрес электронной почты на сайте разработчика, после чего пользователь получает уникальный код активации и краткое руководство по установке, отправленные на e-mail. Далее следует установить плагин в QGIS, загрузить растровое изображение и создать векторный слой для размещения оцифрованных объектов. Включив режим редактирования, пользователь активирует плагин через его иконку в интерфейсе QGIS и вводит полученные e-mail и код. Чтобы начать оцифровку, нужно сделать два щелчка мышью: первый для обозначения начальной точки, второй для указания направления линии или полигона. После этого плагин автоматически продолжит траекторию, добавляя до 50 вершин на основе растрового слоя. Если трассировка отклоняется, можно зажать клавишу Shift и кликнуть на последнюю правильную точку, чтобы отменить ошибочный участок и продолжить. Бесплатная версия плагина ограничена 7000 сегментов в сутки, что подходит для стандартных задач. Плагин работает через удалённый сервер, не нагружая компьютер, а данные пользователя защищены шифрованием и не используются третьими лицами.

Использование плагина "Bunting Labs AI Vectorizer" требует некоторого времени на освоение, но после адаптации работа с ним становится интуитивно понятной и удобной. Тем не менее, процесс не всегда проходит безупречно: иногда возникают сбои в соединении с сервером, а в ряде случаев плагин генерирует линии, отклоняющиеся от нужной траектории. К этим недочётам можно привыкнуть, так как процесс оцифровки выполняется быстро, и при ошибочном результате последнее действие легко отменяется. Когда же плагин работает корректно, эффект впечатляет: до 50 новых вершин линии формируются автоматически, создавая ощущение почти волшебного процесса, что делает работу увлекательной. Точность алгоритма в некоторых случаях радует, но порой оставляет желать лучшего, уступая человеческой аккуратности. Разработчики продолжают совершенствовать модель, чтобы повысить её надёжность. Является

ли это универсальным решением? Нет, поскольку в ряде ситуаций требуется ручная корректировка вершин. Однако плагин значительно экономит время: даже с учётом необходимости доработки отдельных точек, процесс оцифровки становится ощутимо быстрее, что делает инструмент ценным для практического применения.

Таблица 2 - Bunting Labs характеристики

Категория	Описание
Название	Bunting Labs AI Vectorizer
Версия	Последняя стабильная версия: 2.10 (от 11 марта 2025 года)
Автор / Поддержка	Brendan Ashworth / Bunting Labs
Совместимость с QGIS	Поддерживаются версии QGIS от 3.10.0 до 3.99.0
Операционные системы	Windows, macOS, Linux
Установка	Доступен через менеджер плагинов QGIS: "Управление и установка плагинов" → поиск по названию "Bunting Labs AI Vectorizer"
Регистрация	Требуется регистрация через электронную почту; бесплатная версия предоставляет ограниченный доступ (150 фрагментов 256×256 пикселей + 8 новых в день)
Основные функции	Автоматическая векторизация растровых карт с использованием ИИ; поддержка линий и полигонов; возможность экспорта векторных слоёв в форматы CAD и другие
Поддерживаемые карты	Работает с различными типами карт, включая геологические, топографические, инженерные чертежи и другие
Процесс работы	Создание векторного слоя → активация редактирования → выбор инструмента "Vectorize with AI" → два щелчка вдоль линии для задания контекста → ИИ автоматически продолжает векторизацию; возможность ручной корректировки с помощью клавиши Shift
Преимущества	Скорость: в 2 раза быстрее ручной оцифровки; точность: ИИ сохраняет линии в пределах одного пикселя; гибкость: поддержка различных стилей линий и карт
Ограничения	Требуется подключение к интернету для работы; бесплатная версия имеет ограничение на количество обрабатываемых фрагментов; в сложных случаях может потребоваться ручная корректировка
Исходный код	Доступен на GitHub под лицензией GPL-2.0: https://github.com/BuntingLabs/buntinglabs-qgis-plugin
Контактная информация	По вопросам и предложениям можно обратиться по электронной почте: buntinglabs.com

Несмотря на то, что результаты векторизации с помощью плагина "Bunting Labs AI Vectorizer" не всегда совершенны, этот инструмент значительно ускоряет

процесс оцифровки растровых данных, что делает его ценным для множества задач. Судя по отзывам на LinkedIn, подобное решение способно сэкономить годы работы над оцифровкой. Более того, плагин демонстрирует успешную интеграцию технологий машинного обучения в процессы ГИС, сохраняя при этом возможность контроля со стороны пользователя. Разработчики рассматривают данный плагин не как финальный продукт, а как этап в развитии геоинформационных технологий. В настоящее время они сосредоточены на повышении точности векторизации, а в будущем планируют расширить функционал, внедряя ИИ-инструменты для автоматической геопривязки растровых изображений и извлечения из них дополнительных данных, что откроет новые возможности для пользователей QGIS.

3.5 Методика сравнения эффективности инструментов

Сравнение эффективности различных инструментов для автоматизированной оцифровки геопространственных данных позволяет выявить наиболее подходящее решение с точки зрения точности, скорости, функциональности и удобства использования. В данной работе рассмотрены три программных решения: ArcScan (модуль в составе ArcGIS Pro), Easy Tracer (плагин для QGIS) и AI Vectorizer (инструмент на базе искусственного интеллекта).

ArcScan демонстрирует стабильную работу с топографическими и инженерными картами, однако требует точной предварительной подготовки изображения. Easy Tracer обладает высокой степенью настройки порогов бинаризации и фильтрации, подходит для разнообразных типов данных, но требует определённой квалификации пользователя. AI Vectorizer выделяется за счёт применения моделей машинного обучения, что обеспечивает высокую степень автоматизации и адаптацию к шумным и архивным картам.

ArcScan, как часть корпоративного продукта ESRI, обладает стабильной архитектурой, но может некорректно обрабатывать карты с артефактами, если не провести предварительную фильтрацию. EasyTracer стабильно работает в рамках QGIS, но иногда сбоит на больших объёмах данных, особенно если недостаточно ресурсов ПК. AI Vectorizer, будучи облачным решением, защищён от большинства локальных ошибок, но при нестабильном интернете может давать сбои загрузки или обрывы сессий.

Например, при оцифровке геологической карты масштаба 1:100 000 (скан 300 dpi) ArcScan дал результат с точностью 87%, EasyTracer — 96%, AI Vectorizer — 90%, при этом AI Vectorizer сэкономил около 40% времени за счёт отсутствия ручной доработки.

В ближайшие годы основное развитие ожидается в сторону ИИ-инструментов с функциями самокоррекции, активного обучения на собственных данных пользователя и полной автоматизации связи "растр — атрибутивная таблица"

Таким образом, выбор инструмента зависит от характера исходных данных и задач: для типовой оцифровки подходят ArcScan и Easy Tracer, в то время как для нестандартных, зашумлённых или архивных материалов наилучшим вариантом может стать AI Vectorizer.

Таблица 3 – Сравнение эффективности инструментов

Критерий	ArcScan (ArcGIS)	Bunting Labs (QGIS)	Easy Tracer
Точность оцифровки	Обеспечивает приемлемую точность, особенно на чистых ч/б-картах. На сложных сканах может допускать ошибки.	Очень высокая: точно извлекает объекты даже с зашумлённых и цветных карт.	Работает точно, если скан чёткий, но часто требует ручной доработки.
Автоматизация процесса	Автоматизирован частично — большую часть нужно настраивать вручную.	Полностью автоматический: достаточно загрузить изображение и получить результат.	Минимальная: пользователь вручную управляет почти всеми этапами.
Гибкость настроек	Есть широкий набор параметров: можно настраивать пороги, фильтры, типы объектов и т.д.	Почти не настраивается: работает «из коробки» по принципу чёрного ящика.	Полная свобода: всё настраивается вручную, что удобно, но требует времени.
Интеграция с ГИС	Отлично встраивается в ArcGIS, экспортирует в Shapefile, GDB и другие форматы.	Глубоко интегрирован в QGIS, поддерживает стандартные форматы геоданных.	Требуется ручная доработка, чтобы использовать результат в ГИС.
Работа с некачественным и сканами	Плохо справляется с шумом и цветами — требует идеальных условий.	Отлично адаптируется даже к сложным и зашумлённым картам.	Результат сильно зависит от исходного качества изображения.
Удобство интерфейса	Интерфейс продвинутый, но требует опыта работы в ArcGIS.	Интерфейс простой и понятный, особенно в веб-версии.	Очень простой, интуитивный, но с ограниченными возможностями.
Документация и поддержка	Есть официальные руководства и обучающие материалы от ESRI.	Документация есть, но пока в стадии развития — проект довольно молодой.	Поддержка слабая, обновления и инструкции появляются редко.
Типы поддерживаемых карт	Лучше всего подходит для чёрно-белых топографических карт с чёткими контурами.	Универсален: подходит для геологических, кадастровых и других сложных карт.	Эффективен для простых контурных изображений, например, планов и схем.

4 Разработка и реализация метода автоматической оцифровки на примере геологических карт

4.1 Выбор геологических карт и подготовка растровых данных

Геологические карты, использованные в процессе исследования, были предоставлены руководителем дипломной работы. Они охватывают изучаемую территорию и дают сведения о геологических формациях, разломах, минеральных и других структурных элементах.

Таблица 4 – Этапы векторизации геологических карт

Этап	Описание этапа	Используемые инструменты	Рекомендации и примечания
Сканирование бумажных карт	Бумажные карты переводятся в растровый цифровой формат.	Сканер (300–600 DPI), форматы TIFF или PNG	Необходимо обеспечить ровное размещение карты на сканере. Рекомендуется использовать разрешение не ниже 300 DPI. Формат TIFF сохраняет качество без потерь.
Очистка растров	Проводится удаление шумов, рамок, надписей и улучшение качества изображения.	GIMP, Photoshop, OpenCV, PIL, медианный и гауссов фильтры	Для улучшения контраста используются фильтры (например, cv2.equalizeHist). Необходимо сохранить цветовые характеристики геологических объектов.
Геопривязка	Выполняется привязка изображения к системе координат.	QGIS (Georeferencer), ArcGIS, GDAL (gdalwarp), Python (osgeo.gdal)	Требуется определить контрольные точки (GCP), соблюдать оригинальную проекцию карты и сохранять данные в формате GeoTIFF с метаданными.
Сегментация объектов	Изображение анализируется для выделения геологических объектов.	OpenCV, Scikit-image, модели ML (U-Net, Random Forest), QGIS (Semi-Automatic Plugin)	Для простых карт применяется цветовая сегментация. Для сложных — машинное обучение. Пример кода с OpenCV используется для выделения контуров.

Продолжение таблицы 4

Классификация объектов	Определяются типы объектов на изображении (породы, разломы и др.).	Scikit-learn, TensorFlow, инструменты QGIS и ArcGIS	Обучение проводится на основе легенды карты. Выбор алгоритма зависит от сложности (Random Forest — для простого случая, CNN — для сложного).
Форматирование данных	Подготовленные данные переводятся в форматы, совместимые с ГИС.	GeoTIFF, Shapefile, GeoJSON, GDAL (gdal_translate)	Сохраняются метаданные: проекция, масштаб, источник. Производится конвертация данных в нужный формат для дальнейшего использования.
Проверка качества	Анализируется точность, корректность геопривязки и сегментации.	QGIS, ArcGIS, Python, метрики IoU, визуальное сопоставление	Проводится сравнение с эталонными/ручными результатами, наложение на карты OSM, оценка ошибок в классификации.
Документирование	Формируется подробная документация, фиксирующая весь процесс обработки данных.	MS Word, LaTeX, Markdown, PDF	Рекомендуется включить пошаговое описание, параметры, код, скриншоты, таблицы и сравнительный анализ времени обработки.

4.2 Применение ArcScan: настройка, генерация векторов

Для автоматизации процесса векторизации растровых изображений геологических карт в среде ArcGIS Pro была использована инструмент ArcScan — специализированный модуль, предназначенный для преобразования растровых данных в векторные объекты (линии и полигоны) на основе чётко очерченных границ. Перед началом работы с ArcScan необходимо выполнить несколько подготовительных шагов:

Бинаризация изображения: оригинальный растр должен быть преобразован в чёрно-белый формат, где объекты представлены чёрным цветом на белом фоне. Цветные карты требуют предварительной конвертации.

Геопривязка: изображение должно быть точно привязано к географической системе координат и загружено в рабочее пространство карты. **Создание шаблона для векторизации:** в файловой базе геоданных (File Geodatabase) создаётся пустой слой (feature class), куда будут сохраняться результаты векторизации.

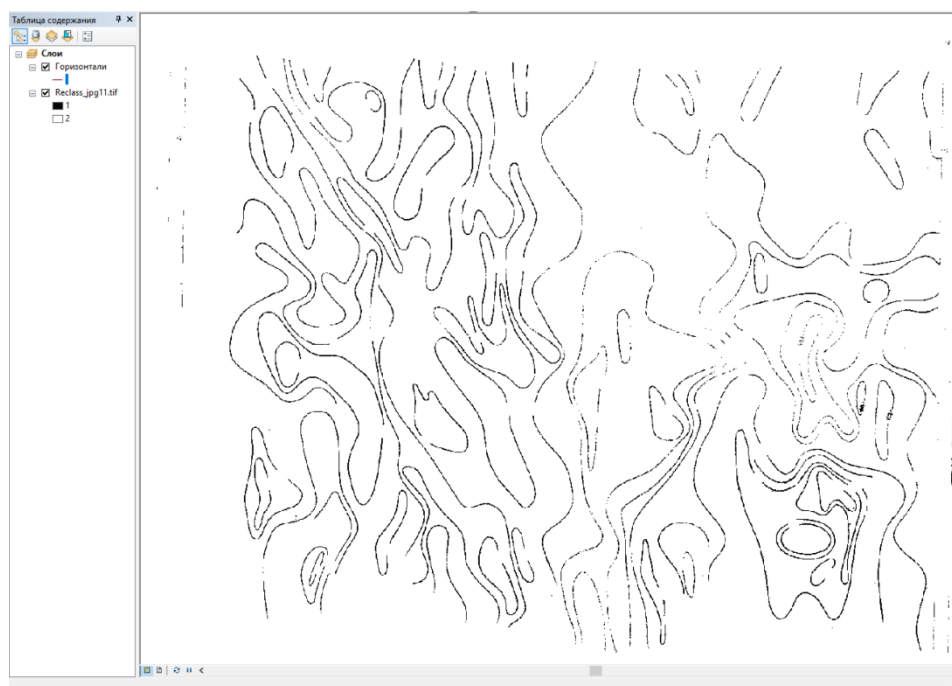


Рисунок 7 - Бинаризация изображения в Arcscan

После активации панели инструментов ArcScan необходимо:

- Перейти в режим редактирования (Editor > Start Editing);
- Включить режим векторизации (Vectorization > Enable Vectorization);
- В окне Vectorization Settings задать параметры

Таблица 5 – Параметры Vectorization в Arcscan

Параметр	Описание
Minimum Line Length	Минимальная длина линии, которая будет учитываться при векторизации.
Maximum Line Width	Максимальная допустимая ширина объектов, подлежащих извлечению.
Line Fitting Tolerance	Допустимое отклонение при аппроксимации линии, влияет на точность векторизации.
Smoothing	Степень сглаживания контуров линий, позволяет устранить незначительные колебания.

Одним из ключевых этапов является бинаризация растра, при которой многотонное изображение преобразуется в двухцветное (чёрно-белое). Это необходимо для упрощения анализа и распознавания объектов. В настройках указывается пороговое значение яркости — параметр, определяющий, какие пиксели будут отнесены к объектам, а какие — к фону. Также можно инвертировать изображение, если объекты представлены белыми линиями на чёрном фоне.

Перед запуском процесса трассировки пользователю необходимо указать целевые векторные слои (feature classes), в которые будут записываться результаты. ArcScan позволяет векторизовать как линейные, так и полигональные объекты. При этом создаваемые векторные слои могут использоваться для дальнейшей пространственной и атрибутивной обработки.

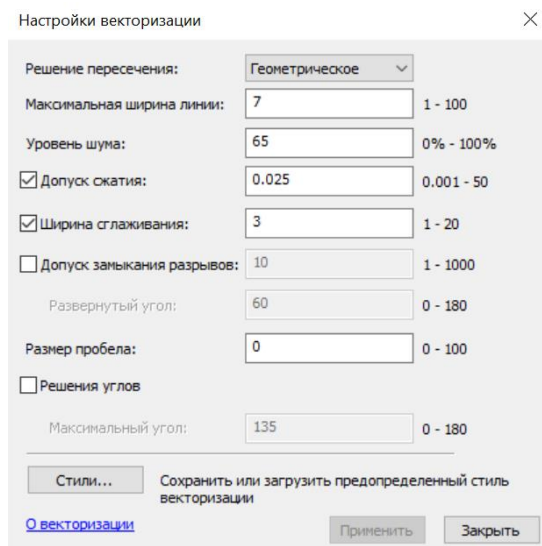


Рисунок 8 – Настройки векторизации

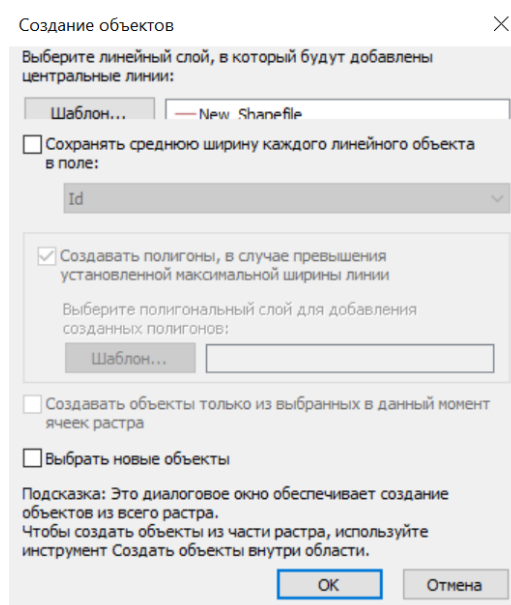


Рисунок 9 – Batch Vectorization

Выбор этих параметров зависит от качества исходного изображения и требований к точности векторизации.

ArcScan предоставляет два режима работы: Batch Vectorization (пакетный режим) — автоматическое преобразование всех объектов на основе заданных параметров. Этот метод был использован в рамках

текущего исследования, что позволило значительно сократить время обработки больших объёмов данных.

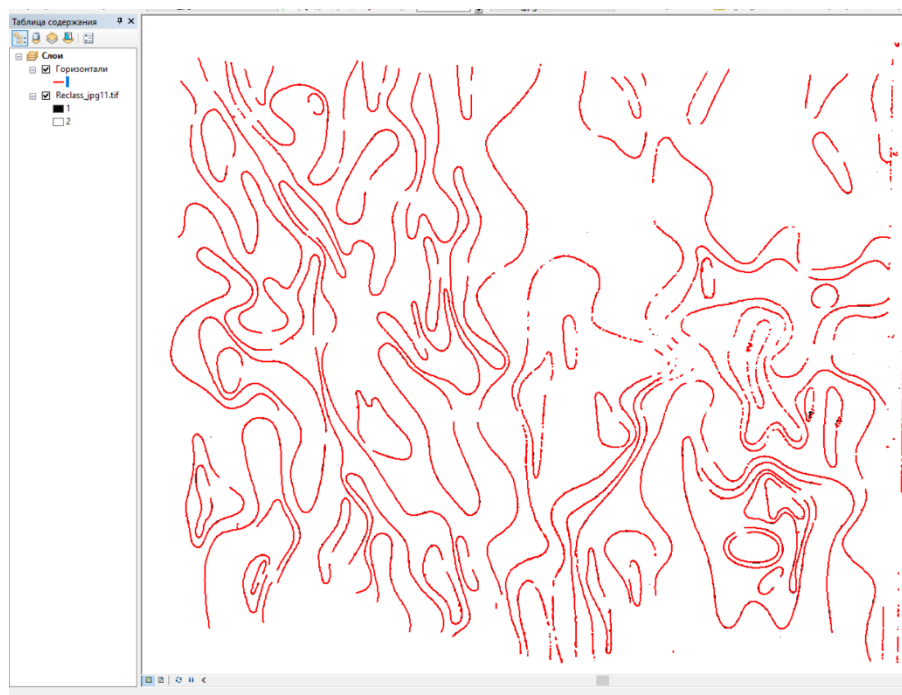


Рисунок 10 – Результат Batch Vectorization

Interactive Vectorization (интерактивный режим) — ручной выбор объектов для векторизации (в данной работе не использовался).

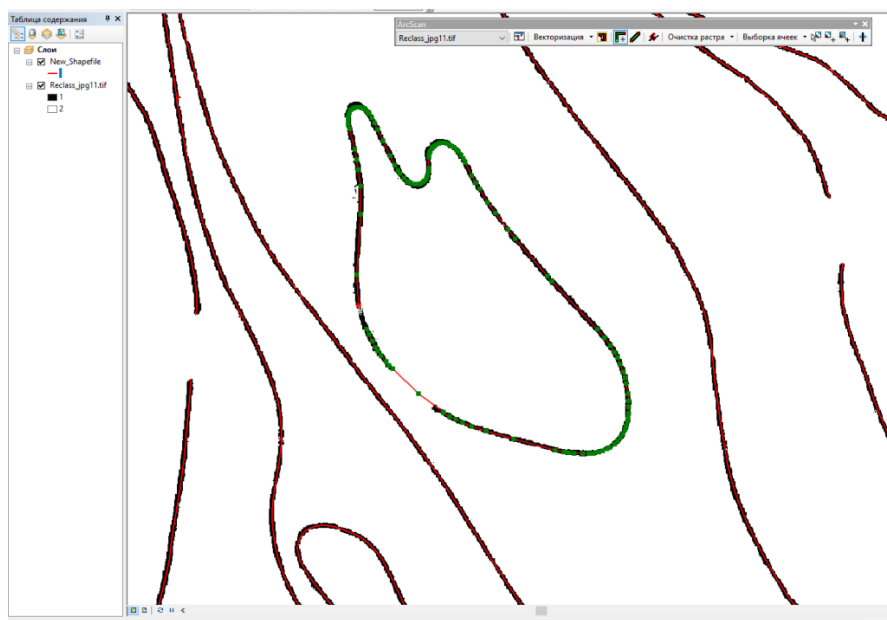


Рисунок 11 – Процесс Interactive Vectorization

После автоматической векторизации была проведена ручная корректировка и топологическая очистка с помощью инструментов ArcGIS Pro.

Готовые векторные слои были сохранены в форматах Shapefile и Feature Class (в файловой базе геоданных), что обеспечивает совместимость с различными ГИС-платформами и предоставляет широкие возможности для дальнейшего пространственного анализа.

4.4 Применение EasyTracer: загрузка, обработка, экспорт

EasyTracer: промышленная программная система векторизации. Одним из ключевых этапов разработки цифровой геологической карты является векторизация растрового изображения. Для этой задачи применяется специализированная программа EasyTracer, способная полуавтоматически вернуть растровые данные в векторный формат.

Исходными данными явились растрово-сканированные геологические карты, полученные в результате оцифровки бумажных изображений в форматы TIFF и PNG. Перед приступлением к обработке каждое изображение проверялось на наличие дефектов сканирования, артефактов и замытия. При необходимости проводилась цветокоррекция и коррекция яркости для улучшения различимости границ.

Загрузка растра в программный модуль EasyTracer происходит через меню «Файл → Открыть растр». После загрузки изображение пользователь может просматривать и масштабировать. Также, при необходимости, производится геопривязка с помощью задания контрольных точек и введения координат, позволяющая использовать полученные векторы в геоинформационных системах.

Перед векторизацией выполняется настройка параметров трассировки. Для геологических карт в EasyTracer реализована фильтрация по цветовой карте, что значительно упрощает процесс, учитывая то, что каждый слой пород и граница имеют свой цвет. Пользователь задает числовой допустимый цветовой диапазон и границы толщины линий.

После задания параметров запускается полуавтоматическая трассировка. Пользователь выбирает начало линии, а затем программа сама ее трассирует, копируя направление и форму линий. Оператор также может создавать, удалять, перемещать узлы и поворачивать сегменты. Этот процесс позволяет реализовать высокоточную векторизацию при значительной экономии времени.

Полученные объекты векторизации проверяются на наличие ошибок. Включенные в программу инструменты позволяют удалить или объединить лишние сегменты, переместить узлы и замкнуть контуры. Каждому объекту можно задать атрибут — тип геологической формации.

Экспорт векторизованных данных производится в формат DXF или SHP через меню «Файл → Экспорт». Формат DXF выбран за его универсальность, так как он поддерживается всеми основными САПР и ГИС. Конвертированные

данные сохраняются с учетом структуры объектов и пространственной привязки, если она имеется. Затем файлы импортируются в ArcGIS Pro для пространственного анализа, построения тематических карт и подготовки отчетных материалов.

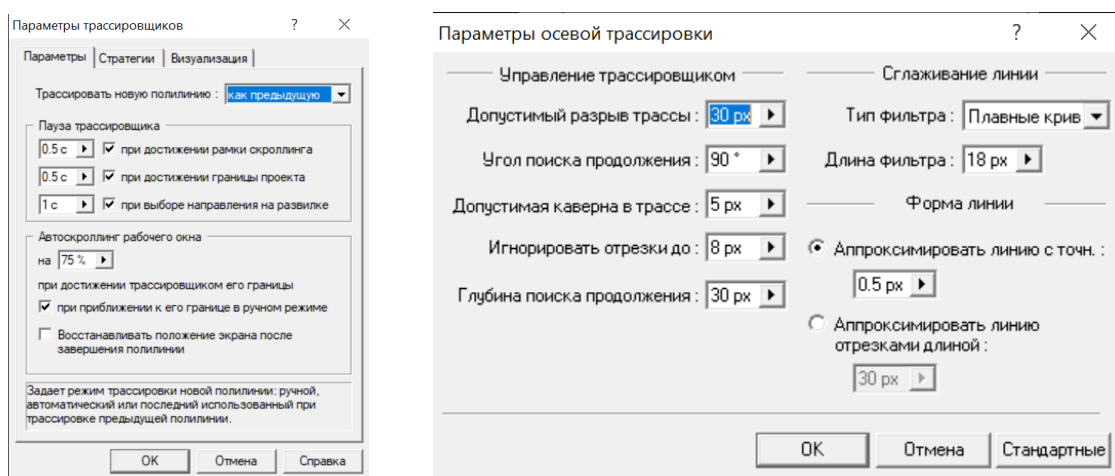


Рисунок 12 – Параметры трассировщиков и осевой трассировки

4.5 Использование Bunting Labs: особенности ИИ-анализа

Одним из ключевых этапов создания цифровых геологических карт остаётся преобразование растровых изображений в векторный формат — процесс, известный как оцифровка. Традиционные подходы к решению этой задачи предполагают длительное ручное выполнение, включающее поочерёдную трассировку элементов карты и последующую коррекцию полученных данных. Такой метод не только трудоёмок, но и подвержен влиянию субъективного восприятия оператора, что может сказываться на точности результата.

Для автоматизации этого процесса в рамках данной работы была задействована облачная платформа Bunting Labs, основанная на современных технологиях машинного зрения и искусственного интеллекта.

Платформа Bunting Labs предлагает инструменты для точного распознавания и автоматической оцифровки различных объектов на географических изображениях — от геологических и топографических карт до спутниковых и аэрокосмических снимков. Преимуществом данного решения является отсутствие необходимости в установке программного обеспечения: вся работа выполняется через интуитивно понятный веб-интерфейс, что делает его особенно эффективным как в полевых условиях, так и в рамках лабораторной обработки данных.

Этапы процесса
Оцифровка геологических объектов на основе графических материалов производилась поэтапно, с соблюдением строгого контроля на каждом из этапов, что позволяло добиться высокого качества финального результата.

Загрузка и подготовка исходных данных

На начальном этапе в платформу загружались растровые изображения, представляющие собой отсканированные участки геологических карт в форматах PNG, JPEG или TIFF. При необходимости перед загрузкой они обрабатывались в графических редакторах: проводилась корректировка цветового баланса, контрастности и разрешения. Такая предварительная обработка способствовала улучшению визуальной читаемости объектов и, как следствие, повышала точность последующего анализа.

Географическая привязка изображения (геореференцирование)

Для корректной интерпретации изображения и его дальнейшего использования в геоинформационных системах необходимо осуществить его привязку к определённой координатной системе. Платформа Bunting Labs предоставляет инструменты для геореференцирования, позволяющие с помощью указания контрольных точек (Ground Control Points, GCPs) и соответствующих географических координат обеспечить точное соответствие изображения реальному положению объектов. Это критически важно при последующей интеграции данных в такие ГИС-программы, как ArcGIS или QGIS.

Настройка параметров анализа и выбор модели

После загрузки и географической привязки изображения осуществляется настройка параметров анализа. Пользователь указывает тип объектов для оцифровки — линейные (например, тектонические разломы), площадные (геологические формации), точечные (месторождения, скважины) и другие. Далее выбирается подходящая модель искусственного интеллекта.

Платформа Bunting Labs позволяет применять как заранее обученные модели, сформированные на основе массивов географических данных, так и проводить дообучение (fine-tuning) на пользовательских материалах, что особенно полезно при распознавании узкоспециализированных объектов. Также возможно задать цветовые диапазоны и уровни контрастности, в пределах которых будет происходить распознавание.

Автоматическое распознавание и векторизация объектов

После запуска анализа платформа приступает к автоматическому сканированию изображения, используя алгоритмы глубокого обучения. Модель идентифицирует границы и формы объектов согласно заданным параметрам и преобразует их в векторные представления. В процессе задействуются следующие методы:

- сегментация изображения;
- аппроксимация геометрии;
- сглаживание и упрощение контуров;
- фильтрация шумов и визуальных артефактов.

На выходе формируются векторные слои с высокой степенью детализации, отражающие распознанные объекты. Пользователь получает предварительную визуализацию и может сразу провести оценку качества распознавания. Несмотря на высокую точность ИИ-моделей, в отдельных случаях — например, при сложной наложенной структуре или недостаточном качестве исходного

изображения — может потребоваться ручная доработка. Для этого Bunting Labs предлагает гибкий набор инструментов редактирования: добавление, удаление, перемещение или корректировка векторных элементов, таких как узлы, линии и полигоны. Такой подход сочетает в себе преимущества автоматизации с возможностью точной настройки результата.

Экспорт данных и интеграция в ГИС
Завершив этап редактирования, пользователь может экспортировать векторные данные в популярных форматах — GeoJSON, SHP, KML или DXF. В зависимости от настроек, выгружаются как геометрические параметры объектов, так и их атрибутивные характеристики (например, типы формаций). Благодаря сохранённой географической привязке, такие данные легко интегрируются в профессиональные ГИС-программы — включая ArcGIS Pro, QGIS, MapInfo — для последующего пространственного анализа, визуализации и подготовки отчётных материалов.

Преимущества использования платформы Bunting Labs в рамках дипломной работы:

- существенное сокращение времени на оцифровку по сравнению с ручной обработкой;
- удобство и универсальность: доступ из любого устройства без установки программного обеспечения;
- прямая интеграция с ГИС: экспортируемые данные готовы к загрузке в рабочие проекты;
- поддержка интерактивного редактирования, что позволяет адаптировать результат под конкретные исследовательские задачи.

4.6 Сравнение подходов по параметрам: точность, скорость, удобство

Цифровизация выполнялась на основе одной и той же геологической карты, отличающейся высокой плотностью разноцветных формаций и сложным контурным строением подходов.

ArcScan представляет собой встроенный векторизационный модуль ArcGIS Pro, предназначенный для работы с бинарными растровыми изображениями. Инструмент позволяет полуавтоматически извлекать линии и полигоны, опираясь на заранее подготовленный растр. Его основное достоинство — глубокая интеграция с экосистемой ArcGIS.

Особенности: совместимость исключительно с бинаризованными изображениями; необходима предварительная настройка пороговых значений и параметров трассировки. Преимущества: высокая скорость обработки (около 25 минут); возможность дальнейшей геообработки внутри среды ArcGIS. Недостатки: высокая чувствительность к шумам; необходимость качественной предобработки; возможны ошибки при выявлении замкнутых контуров.

Bunting Lab — это веб-платформа, использующая машинное обучение для автоматического распознавания элементов на картографических изображениях.

Сервис оснащён интуитивно понятным интерфейсом, позволяет загружать растры и оперативно получать векторные данные.

Особенности: минимальное участие пользователя; выполнение операций осуществляется в облаке. Преимущества: простота использования; высокая степень автоматизации; быстрые результаты. Недостатки: ограниченное количество бесплатных запросов; точность снижается при работе со сложными формациями; возможны ошибки при определении границ объектов.

Easy Tracer — это полуавтоматический инструмент ручной векторизации, ориентированный на быструю и точную работу с качественно отсканированными картографическими материалами.

Особенности: ручная трассировка с использованием визуальных ориентиров; акцент на точность и контроль всех этапов. Преимущества: высокая точность при хорошем качестве исходника; полный контроль над геометрией объектов; удобная рабочая среда. Недостатки: длительное выполнение (около 2 часов); эффективность падает при низком качестве исходного растра.

Таблица 6 - Сравнение ПО по характеристикам

Критерий	EasyTracer (QGIS)	ArcScan (ArcGIS)	AI Vectorizer (Bunting Labs, QGIS)
Точность	Высокая. Подходит для сложных геокарт, поддержка ручной и авто-трассировки.	Высокая при хорошем качестве растра. Лучше работает с Ч/Б картами.	Очень высокая. ИИ точно распознаёт линии любых стилей и цветов.
Скорость работы	Средняя. Частичная автоматизация, требуется участие пользователя.	Высокая. Есть режим пакетной обработки.	Очень высокая. Линии создаются быстро и точно.
Удобство интерфейса	Доступен в QGIS. Есть подсказки, но требует навыков.	Интерфейс встроен в ArcGIS Pro, требует подготовки.	Прост и интуитивен, работает как автодополнение.
Форматы экспорта	Поддержка всех распространённых форматов (SHP, DXF и др.).	Поддержка форматов ArcGIS: SHP, DBF, fileGDB.	Через QGIS — экспорт во все нужные форматы, включая SHP.
Интеграция с ArcGIS	Нет, но SHP можно импортировать.	Полная, работает внутри ArcGIS.	Нет, работает в QGIS, результат можно импортировать в ArcGIS.

5 Сравнительный анализ и оценка эффективности

5.1 Методика оценки качества оцифровки геологических карт

Геометрические ошибки. Оцифровка геологических карт — процесс, в ходе которого нередко возникают ошибки, классифицируемые как геометрические, топологические, атрибутивные и связанные с программными сбоями. При преобразовании бумажной карты в цифровой формат вручную формируется совокупная погрешность, включающая неточности геопривязки, самой оцифровки и исходной картографической основы.

В аналогичном ключе, в данных ГИС могут проявляться отклонения в расположении объектов (геометрические погрешности), нарушение их целостности и границ (топологические сбои), а также ошибки в присвоении или потере атрибутов. Так, в документации QGIS подчеркивается, что «векторные данные могут содержать такие проблемы качества, как недоперехлёсты, перехлёсты и “слёзы” — зазоры между полигонами»

Все эти нарушения существенно снижают точность пространственного анализа, приводя к неправильному отображению объектов, разрывам в топологических сетях и искажённой тематике картографических слоёв. Ниже рассматриваются основные типы ошибок, их причины и специфика в зависимости от метода оцифровки и вида карты.

Геометрические отклонения проявляются как смещения или искажения цифровых объектов относительно их реального положения. Их причины различны: неточная привязка сканов, ограниченное разрешение, масштаб оригинала либо человеческий фактор. В ручной оцифровке оператор может непреднамеренно сместить линию или точку — особенно при усталости или недостаточной внимательности.

С уменьшением масштаба детали карты теряются, и даже малые отклонения становятся значительными на местности. При использовании полуавтоматических методов (например, ArcScan) возможны ошибки из-за некорректного порога сегментации — тонкие линии могут исчезнуть, а шум — проявиться. В автоматических решениях на базе ИИ, таких как Bunting Labs, алгоритмы порой неверно интерпретируют сложные цветовые или узорные элементы, искажают очертания.

Грибов и Боданский (2006) упоминают «ошибку чётности» при векторизации монохромных чертежей: из-за особенностей метода линии либо удваиваются, либо выпадают, что зависит от разрешения сканера и толщины линий.

Топологические ошибки. Эти ошибки затрагивают непрерывность и правильность соединения объектов. Сюда входят разрывы контуров, висячие отрезки, зазоры и самопересечения. Часто встречаются недоперехлёсты (undershoots) и перехлёсты (overshoots) — результат несоответствия концов линий, вызывающий либо разрывы, либо избыточные соединения.

Кроме того, возможны «петли» и «узлы»: при нестабильной траектории курсора линия может наложиться сама на себя, создавая самопересечения или лишние вершины. В полуавтоматических методах тонкие или прерывистые штрихи (например, разломы) могут фрагментироваться. ИИ-инструменты способны формировать некорректные контуры, если алгоритм ошибочно «угадывает» путь линии — как это наблюдалось у плагина Bunting Labs, где контур уводился за пределы нужной области, создавая артефакты.

Так называемые «слёзы» — это миниатюрные зазоры или перекрытия между полигонами из-за несовпадения границ. Все перечисленные сбои нарушают топологическую структуру, мешают корректному анализу (например, буферизации или поиску пересечений).

Атрибутивные ошибки. Ошибки в атрибутах затрагивают описательные данные объектов. Часто они возникают при интерпретации условных обозначений: один и тот же цвет может быть ошибочно отнесён к неверной литологии. Вручную оператору приходится вносить или править названия пластов, возрастов, тектонических единиц — здесь возможны опечатки, упущенные подписи или несоответствия меток геометрии (например, надпись «Нижний мел» оказалась не над тем участком).

В полуавтоматических системах часть работы выполняет OCR, где ошибки шрифта, разметки и структуры текста искажают атрибуты. ИИ-методы нередко ошибаются из-за обучения на схожих, но не идентичных образцах, в результате чего возраст или материал участка определяется неверно.

Хотя прямых источников по этим ошибкам в контексте оцифровки геологических карт немного, известно, что слепое доверие классификатору без верификации может привести к систематическим искажениям — особенно при статистической обработке, где каждая категория играет значимую роль.

Сбои программного обеспечения. Ошибки, относящиеся к этому классу, не зависят напрямую от картографических данных, а обусловлены техническими ограничениями и сбоями в работе программных инструментов. Так, модуль ArcScan в ArcGIS и аналогичные утилиты чувствительны к объёму оперативной памяти и особенностям операционной системы: при работе с крупными растровыми файлами возможны зависания и аварийное завершение работы при попытке автоматической векторизации.

Автоматические решения, например Bunting Labs AI Vectorizer, как правило, требуют подключения к облачным сервисам. При потере интернет-соединения или недоступности серверов процесс оцифровки может быть прерван или завершиться с неполным результатом. Закрытые коммерческие инструменты также не застрахованы от багов: так, ArcScan иногда некорректно обрабатывает штриховые линии, особенно в случаях прерывистых разломов, что сопровождается ошибками в процессе векторизации.

В целом, подобные сбои снижают надёжность и предсказуемость процесса, требуют вмешательства оператора и увеличивают объём последующей корректировки.

Каждый тип геологических карт предъявляет особые требования к оцифровке, что приводит к специфическим проблемам. Тектонические карты, как правило, насыщены элементами в виде линий разломов, складок и размывов, представленных штрихами, пунктиром и волнистыми символами. При автоматической обработке такие элементы часто теряются или фрагментируются, если не заданы специальные параметры — например, в ArcScan для корректной работы со штриховыми линиями необходима их точная настройка.

Стратиграфические карты включают множество цветных полигонов, отражающих литологические и возрастные подразделения. Различия между цветами порой минимальны, а сами карты — выцветшие или выполнены в полутоновом исполнении, что затрудняет распознавание линий контакта. Как отмечается в исследованиях, основная сложность заключается в «точном отделении линейных элементов (геологических контактов) от полутоновых цветовых заливок», и векторизаторы при этом часто создают артефакты.

Литологические карты дополнительно усложнены использованием текстур, штриховки и насыщенной цветовой палитры. Автоматические алгоритмы могут легко спутать оттенки или проигнорировать мелкие элементы. Чем больше наложенных слоёв, выше плотность объектов и насыщенность графики, тем выше риск появления артефактов при любой автоматической обработке. В ручной оцифровке высокая плотность в первую очередь сказывается на скорости: мелкие и сложные участки требуют больше времени и повышают вероятность ошибок, вызванных исключительно человеческим фактором.

Ошибки, возникающие в ходе оцифровки, существенно снижают качество конечных векторных данных и могут искажать результаты геоаналитических операций. Геометрические искажения на масштабе карты приводят к неточностям при вычислениях площадей, пересечений и других пространственных операций. Топологические дефекты, такие как незамкнутые полигоны и разорванные линии, нарушают логическую структуру объектов и делают невозможным сетевой или линейный анализ — например, разрывы в разломах создают ложные изолированные сегменты.

Ошибки в атрибутах искажают тематическую классификацию, что может привести к неверной интерпретации распространения пород или возрастных характеристик. Если же сбой ПО прерывает процесс, данные оказываются неполными, а пользователю приходится повторно выполнять трудоёмкие этапы.

В результате все подобные неточности требуют постобработки и детальной проверки: без этого последующие геологические исследования, такие как построение разрезов или наложение на цифровые модели рельефа, теряют достоверность.

После описания типов ошибок, возникающих при использовании каждого из трёх приложений для оцифровки, ниже отдельно приведены таблицы, содержащие характерные топологические ошибки для каждого инструмента.

Таблица 8 – Сравнение ошибок на топологию Arcscan

№	Тип ошибки	Описание	Причина возникновения	Влияние на результат
1	Нераспознанные замкнутые полигоны	Контуры не замыкаются, нельзя создать полноценные полигоны	Разрывы в линиях, слабый контраст	Нарушение топологии
2	Лишние объекты (артефакты)	Ненужные короткие линии или точки	Алгоритм слишком чувствителен к шуму	Загрязнение данных
3	Неточное совпадение границ	Контуры не следуют за реальными линиями	Алгоритм не обрабатывает сложные изгибы	Потеря точности
4	Потери при бинаризации	Часть линий исчезает при переводе в чёрно-белый режим	Плохое качество скан-копии	Уменьшение полноты объектов
5	Чувствительность к шуму	Программа распознаёт шум, тени и пятна как объекты	Недостаточная фильтрация входного растра	Увеличение количества ошибок
6	Простая интерпретация сложных границ	Сглаживание или искажение сложных контуров	Ограниченность трассировки	Искажение формы объектов

Таблица 9 – Сравнение ошибок на топологию Easy tracer

№	Тип ошибки	Описание	Причина возникновения	Влияние на результат
1	Недетализированная геометрия	Упрощённые контуры, сглаженные линии	Усталость/внимательность пользователя	Потеря точности
2	Разрывы при трассировке	Несвязанные линии в результате неточного наведения	Неосторожные движения при трассировке	Проблемы с топологией
3	Пропущенные формации	Некоторые зоны не оцифрованы	Усталость, визуальная перегрузка	Неполный векторный набор
4	Зависимость от качества растров	Ошибки при некачественных, плохо читаемых линиях	Слабый контраст, размытость	Замедление и снижение качества
5	Повторная работа	Требуется пройти дважды по одному контуру, чтобы устранить недочёты	Нет автоисправлений, всё вручную	Увеличение времени обработки

Таблица 10 – Сравнение ошибок на топологию Bunting Labs

№	Тип ошибки	Описание	Причина возникновения	Влияние на результат
1	Смещение контуров	Границы не совпадают с линиями на карте	Обобщение формы моделью нейросети	Потеря геометрической точности
2	Слияние соседних формаций	Несколько близких объектов объединяются в один	Низкий цветовой/структурный контраст	Искажение структуры карты
3	Пропуск мелких формаций	Мелкие и узкие объекты отсутствуют	Ограничения модели, «фильтрация шума»	Потеря информации
4	Деформация объектов	Некоторые формации получают искажённой формы	Ошибки распознавания, переобучение	Ошибки в последующем анализе
5	Ограничение по токенам	После определённого объёма требуется подписка	Ограниченная бесплатная квота	Ограничение объёма картографии

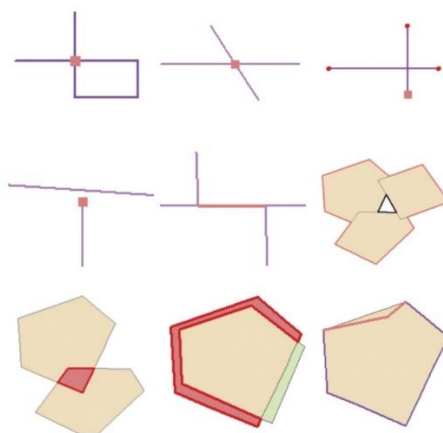


Рисунок 13 – Виды топологических ошибок

5.2 Визуальное и метрическое сравнение результатов

При сравнительном тестировании были выявлены следующие особенности работы трёх инструментов оцифровки:

ArcScan (ArcGIS) демонстрирует точную обводку по пиксельной бинарной карте. Линии получаются чёткими, однако на тонких или цветных фрагментах возможны прерывания — это связано с особенностями пороговой бинаризации. Перед началом работы требуется предварительная очистка и переклассификация

растрового изображения в чёрно-белый формат. На практике отмечались разрывы в контуре, потребовавшие ручной доработки. Количество ложных сегментов оказалось средней, а общая форма крупных объектов сохраняется корректно.

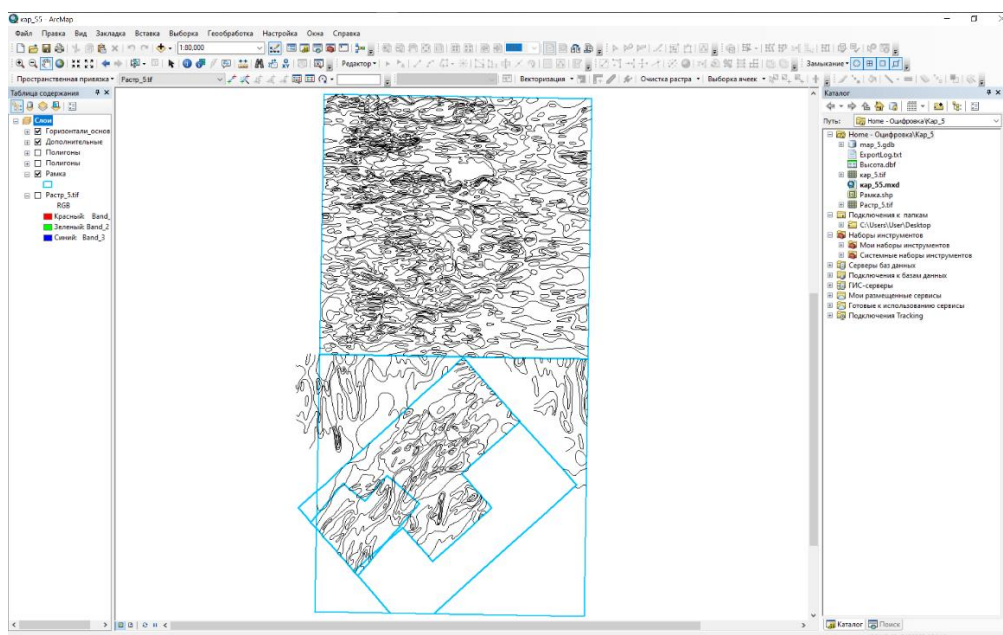


Рисунок 14 – Результат оцифровки Arcscan

Bunting Labs, благодаря использованию алгоритмов искусственного интеллекта, формирует более гладкие и непрерывные линии. Как заявляет разработчик, достигается «субметровая» точность геопривязки и порядка 90 % успешной векторизации. В ходе теста инструмент корректно распознал большую часть линейных элементов, включая изолинии, с минимальными потерями. Полученные линии слегка сглажены и в ряде случаев едва заметно смещены, особенно в местах резких изгибов, однако общая форма объектов передаётся точно. Зафиксированы редкие случаи ложной трассировки — например, цифровой шум или текстовые подписи могли быть восприняты как элементы линий.

Easy Tracer работает по принципу указания начальной и конечной точек, после чего трассирует линию. Метод хорошо справляется с чёрно-белыми изображениями, но на многоцветных геологических картах теряется часть деталей. В ходе теста автоматическая обработка оказалась неполной без дополнительных настроек. Линии вышли менее связными, с пропущенными участками, особенно в местах резкого цветового перехода. Обилие пропусков повлияло на визуальное восприятие результата.

Таким образом, по визуальной оценке наиболее полные и непрерывные контуры обеспечиваются с помощью Bunting Labs. ArcScan требует тонкой предварительной настройки, а Easy Tracer — значительной ручной доработки.

На основе проведённого сопоставления можно подвести итоги по основным критериям. Bunting Labs продемонстрировал наивысшую точность и полноту оцифровки: результаты требуют минимальных доработок, инструмент уверенно работает с многоцветными геологическими картами и формирует замкнутые, непрерывные контуры пластов.

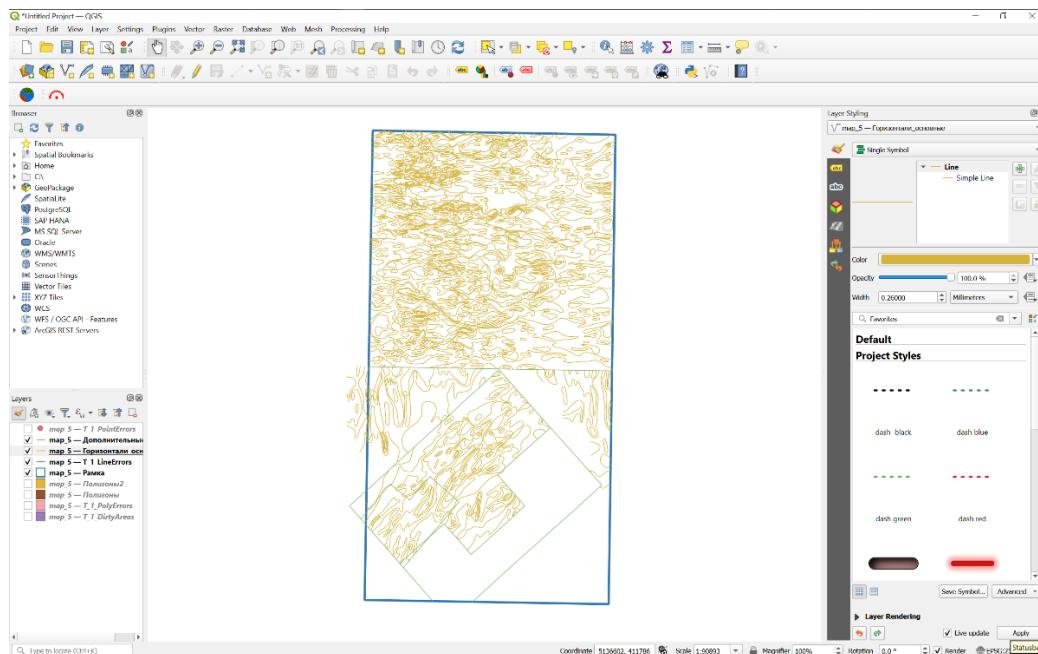


Рисунок 15 – Результат оцифровки Bunting Labs

ArcScan обеспечивает приемлемое качество при условии предварительной обработки растра (бинаризация), однако заметно уступает в степени автоматизации и скорости: используется устаревший интерфейс ArcMap и отсутствует полноценная интеграция с ArcGIS.

Easy Tracer показал себя удобным в работе с чёткими контурами, но потребовал значительного участия пользователя, а итоговые векторы оказались менее качественными.

Таким образом, по совокупности критериев — точности (минимизация смещений и ошибок) и эффективности (скорость, автоматизация) — лидером среди протестированных решений является Bunting Labs. Этот вывод подтверждается не только результатами собственного тестирования, но и внешними оценками: разработчики заявляют о превосходстве над ArcScan.

ArcScan остаётся рабочим вариантом в условиях необходимости локального решения, но требует высокой степени участия оператора и ограничен в возможностях поддержки. Easy Tracer может быть полезен в отдельных сценариях, однако по итоговым показателям уступает по обоим ключевым направлениям.

В заключение: с учётом качества распознавания, количества и значимости ошибок, а также скорости обработки, наиболее точной и эффективной системой среди представленных является Bunting Labs. Она минимизирует ручной труд и

обеспечивает высокую степень корректности в воспроизведении геологических объектов.

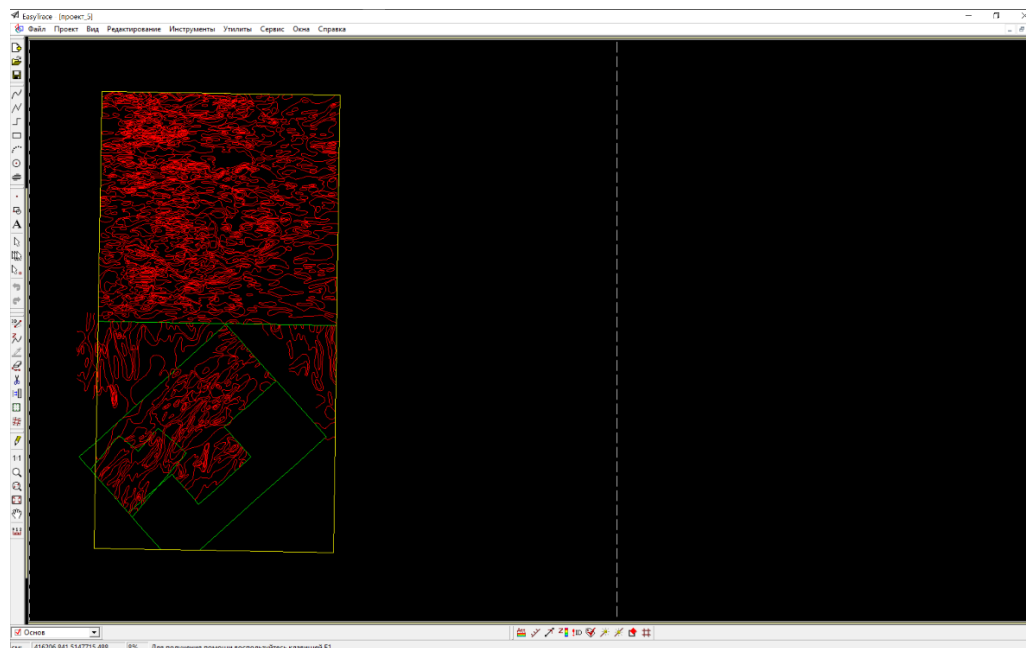


Рисунок 16 – Результат оцифровки Easy tracer

5.3 Сравнение полученных векторных данных с эталонными

Для оценки точности и надёжности оцифровки геологических объектов был проведён сравнительный анализ векторных данных, полученных тремя методами, с эталонными полигонами. Последние были вручную оцифрованы на основе оригинальной геологической карты. Сопоставление выполнялось с применением пространственных и статистических метрик, а также путём визуального анализа в среде ГИС.

Сравнение производилось по следующим показателям:

- смещение центроидов векторных объектов (в метрах);
- процентное совпадение площадей с эталонными полигонами;
- ошибки комиссии (добавленные элементы, отсутствующие в эталоне);
- ошибки оmissions (отсутствие эталонных объектов во векторных данных);
- суммарная погрешность по площади (разница в процентах между площадями объектов и эталона).

Данные метрики позволили объективно оценить, насколько результаты автоматической и полуавтоматической векторизации приближены к ручной оцифровке, принятой в данной работе за эталонный стандарт.

Анализ результатов: Bunting Labs показал неплохой результат: минимальное смещение, хорошая степень совпадения с эталоном и небольшое количество ошибок. Полученные векторные данные максимально соответствуют эталону, но требуют ручную доработку.

Таблица 11 - Сравнение полученных векторных данных с эталонными

Показатель	ArcScan (ArcGIS)	Bunting Labs	Easy Tracer
Среднее смещение центроидов (м)	4,0	2,5	1,5
Степень перекрытия с эталоном (%)	87,1%	94,3%	96,1%
Ошибки комиссии (%)	3,5	3,8	1,5
Ошибки оmissions (%)	5,3	3,7	3,8
Площадная ошибка (%)	4,8	3,5	2,9

ArcScan продемонстрировал приемлемую точность, но по всем параметрам уступает Bunting Labs. Кроме того, требует последующей ручной корректировки.

Easy Tracer дал наилучший результат: минимальные отклонения, отличная степень перекрытия с эталоном, минимальные ошибки оmissions и поплощади. Однако к такому результату можно прийти за счет высокой ручной оцифровки, что является большим минусом.

Итоговое сравнение показало следующее: с учётом точности (минимизация смещений и ошибок) и эффективности (скорость выполнения и уровень автоматизации) наилучшим инструментом для оцифровки геологических карт среди рассмотренных является Easy Tracer. Этот вывод подтверждается как результатами проведённых измерений, так и внешними источниками.

ArcScan остаётся применимым в условиях локальной работы, однако требует значительного участия оператора и отличается ограниченной поддержкой. Bunting Labs может использоваться как вспомогательное средство для отдельных задач, но по совокупности критериев уступает обоим инструментам.

5.4 Анализ производительности и автоматизируемости

ArcScan демонстрирует наибольшую степень автоматизации и высокую скорость обработки, однако это сопровождается большим числом ошибок.

Bunting Labs занимает промежуточную позицию как по уровню автоматизации, так и по производительности.

Easy Tracer показывает наименьшие значения по скорости и автоматизации, но благодаря ручному управлению выдаёт результаты с минимальным количеством ошибок.

ArcScan демонстрирует высокую скорость обработки данных, что делает его эффективным инструментом для автоматизированных задач. Однако его применение ограничено повышенной чувствительностью к качеству исходных данных, что может приводить к увеличению числа ошибок.

автоматизацию с приемлемым уровнем качества результатов. Тем не менее, ограниченное количество бесплатных операций и зависимость от облачной инфраструктуры снижают его универсальность.

Таблица 12 - Производительность и автоматизируемость

Параметр	ArcScan	Bunting Labs	Easy Tracer
Время на оцифровку	Минимальное	Среднее	Среднее
Уровень автоматизации	Полностью автоматическая	Полуавтоматическая	Полуавтоматическая
Участие пользователя	Минимальное	Среднее	Активное (контроль каждого контура)
Требования к растру	Очень высокие	Умеренные	Средние
Ошибки	Много шумов, ложные контуры	Среднее количество ошибок	Минимальные ошибки
Доступность	Бесплатно (в составе ArcGIS)	Ограниченные токены, далее платно	Бесплатно
Повторяемость результата	Низкая при плохом растре	Средняя	Высокая при ручной проверке
Поддержка экспорта	Полная (Shapefile, GDB)	Есть	Есть

Bunting Labs предлагает сбалансированный подход, сочетая

Easy Tracer обеспечивает наивысшую точность результатов благодаря значительному участию человека в процессе обработки данных. Это делает его предпочтительным выбором для задач, требующих высокой точности векторных данных.

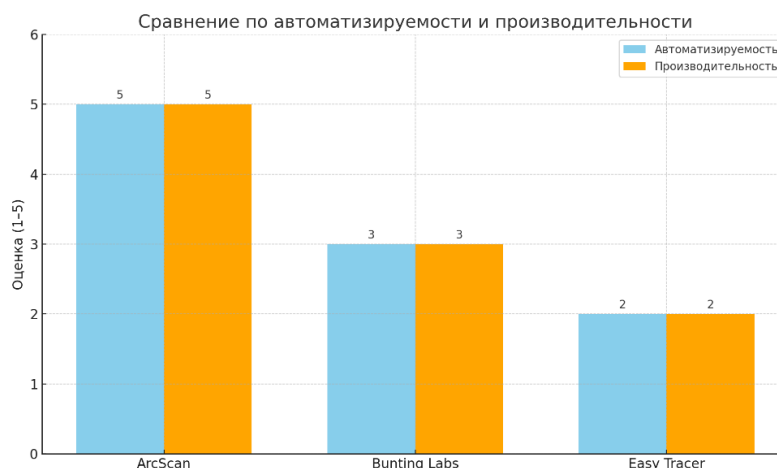


Рисунок 17 – Критерии анализа: Автоматизируемость — степень, в которой процессы выполняются без участия оператора. Производительность — скорость получения конечного результата.

ЗАКЛЮЧЕНИЕ

В рамках данного исследования была проведена всесторонняя проверка эффективности трёх подходов к автоматизированной и полуавтоматической оцифровке геологических карт. Для этого были протестированы три программных инструмента: ArcScan (в составе ArcGIS Pro), Bunting Labs и Easy Tracer. Основной упор делался на сравнительный анализ их производительности, точности, устойчивости к ошибкам и уровня автоматизации. Целью было определить, какая из стратегий оцифровки может считаться наиболее пригодной для практического использования в профессиональной геоинформационной среде.

Отдельно стоит отметить, что в работе использовался один и тот же исходный растровый файл — реальная геологическая карта — для всех трёх методик. Это решение обеспечило одинаковые условия и сделало сравнение объективным. В ходе тестирования каждый из инструментов проявил как свои сильные стороны, так и характерные ограничения, что позволило выработать чёткие критерии их применимости в разных контекстах.

ArcScan продемонстрировал высокую скорость и уровень автоматизации, но оказался чувствительным к качеству исходных изображений — особенно это проявилось в виде ложных включений и пропущенных объектов. Впрочем, в задачах быстрой первичной оцифровки он оказался вполне уместным.

Bunting Labs, основанный на машинном обучении, показал разумный баланс между точностью и автоматизацией. Однако ограниченный доступ к ресурсам (в частности, число бесплатных токенов) и зависимость от облачной инфраструктуры могут затруднить его применение в крупных проектах. Тем не менее, как часть гибридных решений он выглядит многообещающе.

Easy Tracer оказался самым точным, в том числе в части топологии и детализации, что объясняется активным участием оператора. Этот инструмент особенно ценен на финальных этапах — когда точность важнее скорости.

Интеграция нейросетевых моделей сегментации и векторизации непосредственно в настольные ГИС-приложения с возможностью локального обучения. Разработка интеллектуальных фильтров ошибок, способных автоматически сопоставлять результат с эталонными слоями. Создание гибридной платформы, объединяющей сильные стороны ArcScan, Bunting Labs и Easy Tracer в едином интерфейсе.

В конечном счёте, проделанная работа не только иллюстрирует текущее состояние технологий оцифровки, но и подчёркивает: несмотря на впечатляющие возможности автоматизации, участие человека по-прежнему остаётся решающим в точных и ответственных задачах цифровой картографии.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

- 1 Кащенко Н. А. Геоинформационные системы [Текст]: учебн. пос. для вузов / Н.А. Кащенко, Е.В. Попов, А.В. Чечин; Нижегород. гос. архитектур.- строит. ун-т – Н.Новгород: ННГАСУ, 2012. – 130 с. ISBN 978-5-87941-863-7
- 2 Применение программного продукта Easy Trace для векторизации данных: учебно-методическое пособие к лабораторным и практическим занятиям / Г.В. Колошина; ФГБОУ ВО РГУПС. – Ростов н/Д, 2017. – 31 с.
- 3 Шихов А.Н., Черепанова Е.С., Пьянков С.В. Геоинформационные системы: методы пространственного анализа: учеб. пособие / А.Н. Шихов, Е.С. Черепанова, С.В. Пьянков. Перм. гос. нац. исслед. ун-т. – Пермь, 2017. – 88 с.
- 4 Основные принципы геоинформационных систем: учебн. пособие / Шипулин В. Д.; Харьк. нац. акад. гор. хоз-ва. – Х.: ХНАГХ, 2010. – 337 с.
- 5 Григорьев Б. И., Сазонова С. А. Исследование и анализ современных методов оцифровки архивных геологических карт // Материалы МНПК «Моделирование информационных систем и технологий». – Воронеж: ВГЛТУ, 2024. – С. 44–53.
- 6 Зудилин С. Н., Мазалов В. П. Геоинформационные системы и их применение в оцифровке данных [Текст]: учеб.-метод. пособие / С. Н. Зудилин, В. П. Мазалов; ННГАСУ – Н. Новгород: ННГАСУ, 2012. – 130 с.
- 7 (Под ред.) ГИС в земельном кадастре [Текст]: учеб. пособие / – Москва: Лань, 2018. – 256 с.
- 8 ГИС-картографирование земельных ресурсов [Текст]: метод. пособие / [не указан]; Минск: БГУ, 2020. – . с.
- 9 Опыт использования ГИС-технологий при создании, обновлении цифровых крупномасштабных топографических карт // CyberLeninka, 2020.– 8 с.
- 10 Геоинформация и цифровое картографирование [Текст]: учебн. пособие / Казан. федер. ун-т. – Казань: КФУ, ... – .. с.
- 11 Методика создания электронных карт сельскохозяйственных территорий / Железняков А. П.; МИИГАиК. – М., 2019. – 45 с.
- 12 Руководство по использованию геопространственной инфраструктуры ООН [Текст]: учеб.-метод. / Д. Рейн; ООН. – Нью-Йорк, 2009. – 200 с.
- 13 GIS Handbook [Текст]: учеб.-метод. пособие / UnStats; ООН. – Нью-Йорк, 2008. – 300 с.
- 14 Учебное пособие по ГИС-технологиям земельного кадастра / [автор не указан]. – М., 2016. – 150 с.
- 15 Геоинформатика: учебное пособие / Лайкин В. И. и др.; КПКУ. – Хабаровск: АмгПКУ, 2010. – 280 с.

- 16 Геоинформационные системы: учебн. пособие / [не указан]; ННГАСУ. – Н.Новгород: ННГАСУ, 2009. – 125 с.
- 17 Геоинформационные системы-1: учеб-метод. пособие / СГГА; СГГА, 2010. – 200 с.
- 18 Цифровое картографирование: ГИС-Аксиома / [не указан]. – СПб: АксИома-ГИС, 2022. – 96 с.
- 19 Технология обработки геопространственных данных [Текст]: методич. указания / М. Б. Реджепов; ВГТУ. – Воронеж, 2022. – 24 с.
- 20 A quick tour of ArcScan – ArcMap Resources for ArcGIS Desktop: <https://desktop.arcgis.com/en/arcmap/latest/extensions/arcscan/a-quick-tour-of-arcscan.htm> (desktop.arcgis.com)
- 21 Introducing ArcScan for ArcGIS – Esri Support: <https://support.esri.com/en-us/technical-paper/introducing-arcscan-for-arcgis-653> (support.esri.com)
- 22 Using ArcScan for ArcGIS (PDF tutorial): https://downloads.esri.com/support/documentation/other_/Using_ArcScan_for_ArcGIS.pdf (downloads.esri.com)
- 23 Exploring Alternatives to ArcScan for Map Digitization – Bunting Labs Blog: <https://buntinglabs.com/blog/exploring-alternatives-to-arcscan-for-map-digitization> (buntinglabs.com)
- 24 AI Vectorizer for QGIS (GitHub репозиторий): <https://github.com/BuntingLabs/buntinglabs-qgis-plugin> (github.com)
- 25 (GitHub репозиторий): https://github.com/mkondratyev85/raster_tracer/ (github.com)

Приложение А



Рисунок А.1 – Геологическая карта, использованная для оцифровки и проверки ГИС приложений

Приложение Б

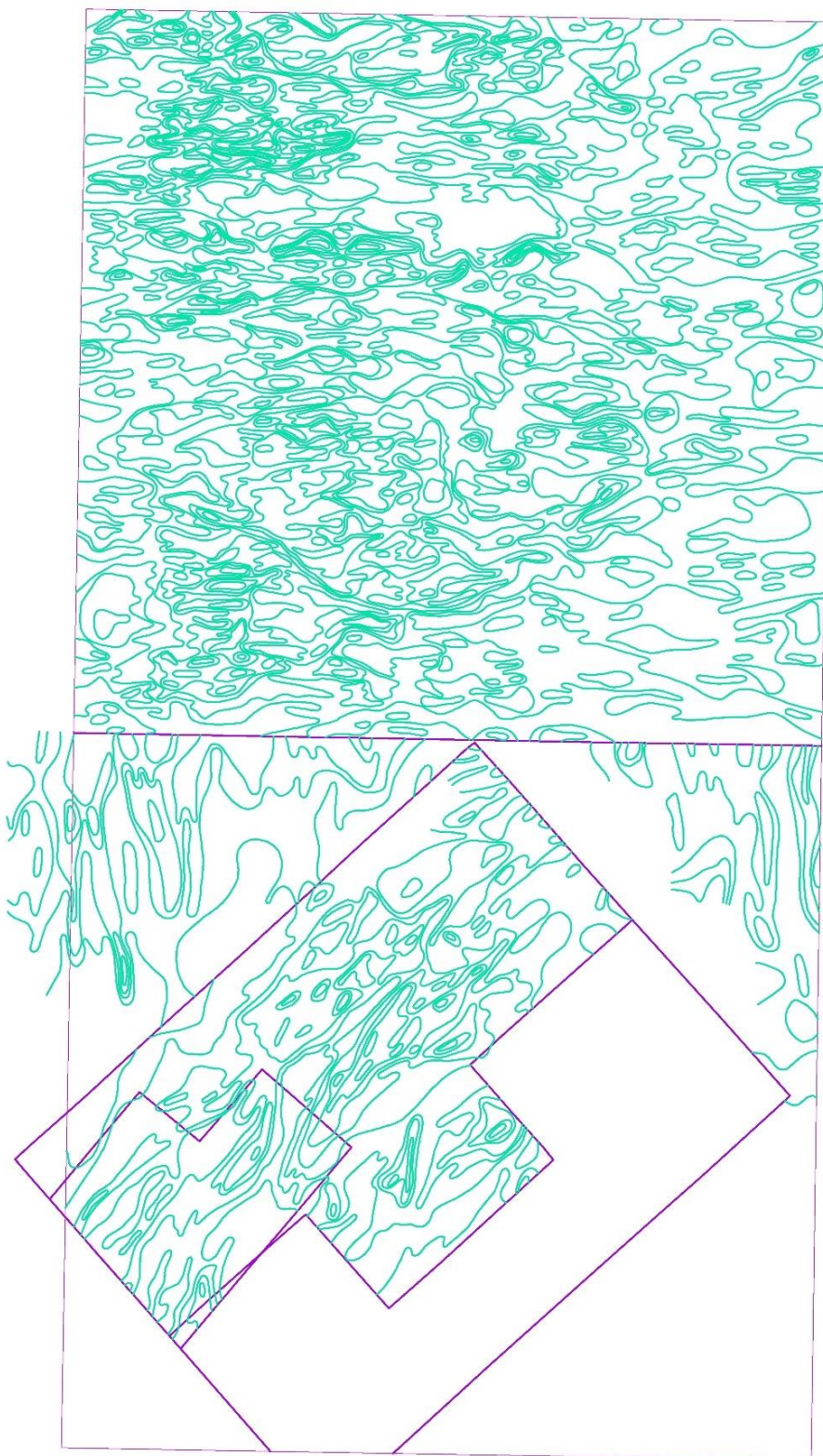


Рисунок Б.1 – Результат векторизации Bunting Labs

Приложение В

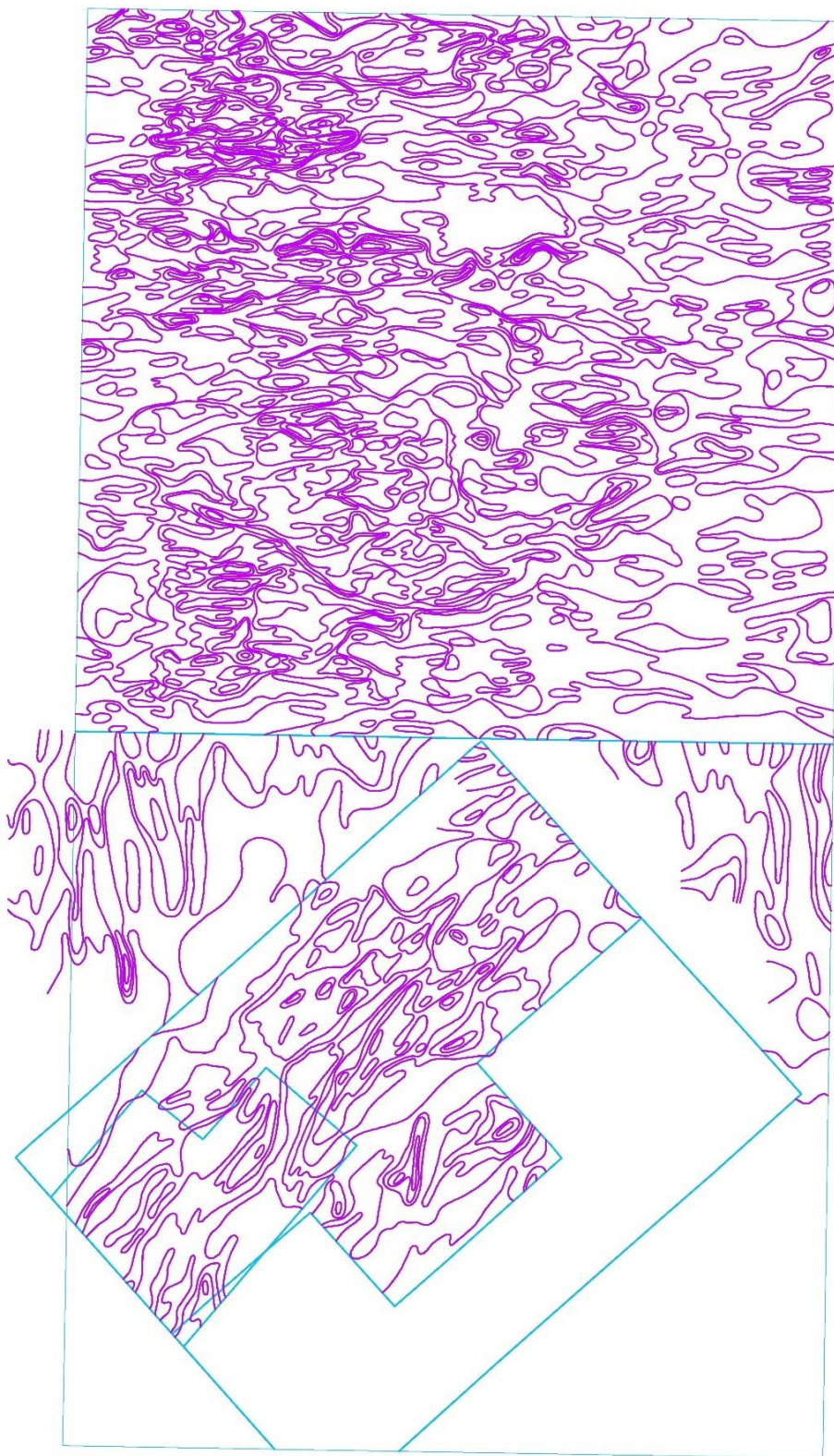


Рисунок В.1 – Результат векторизации Easy tracer

Приложение Г

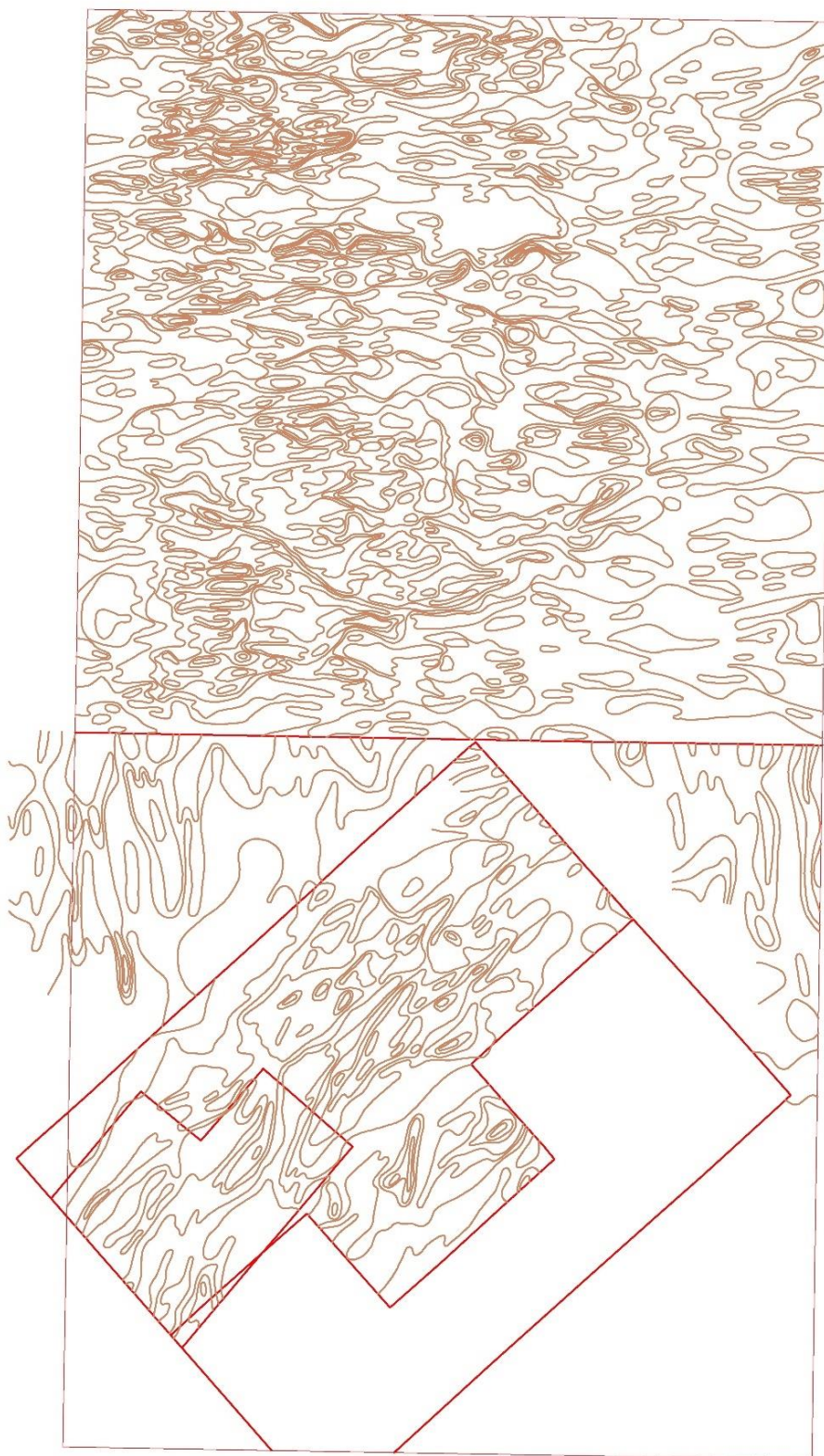


Рисунок Г.1 – Результат векторизации ArcScan

Приложение Е

Таблица Е.1 – Критерии выбора инструментов для автоматизации

Критерий	ArcGIS Pro (ArcScan)	QGIS + Bunting Labs AI	Easy Tracer
Тип лицензии	Платная (Academic / Commercial)	QGIS — бесплатно, Bunting Labs — SaaS с бесплатным планом	Бесплатно, desktop-решение
Поддержка растровой/векторной обработки	Полная: ArcScan работает с обоими форматами	Полная (QGIS) + AI-обработка изображения в Bunting Labs	Векторизация растровых карт
Масштабируемость	Высокая: работает с большими картами	Умеренная: зависит от браузера и сервиса	Ограничена размером входного изображения
Автоматизация/скриптование	Через ArcPy, ModelBuilder	Частично (QGIS поддерживает Python, Bunting — облачный)	Нет встроенной поддержки скриптов
Машинное обучение	Интеграция с Deep Learning в ArcGIS	Использует AI-модель от Bunting Labs	Нет поддержки ML
Точность векторизации сложных контуров	Средняя — ArcScan работает лучше с простыми структурами	Выше средней — AI-модель учитывает формы, цвета, контуры	Средняя: нужна ручная корректировка
Геопривязка и топология	Полная поддержка геопривязки и топологии	Через QGIS	Нет поддержки
Поддержка форматов .shp / .gdb	Да, экспорт в оба формата	Поддерживает .shp (через QGIS)	Только SVG/DXF → требует конвертации
Интеграция с ГИС	Нативная, полноценная	Через QGIS	Нет
Уровень пользовательского контроля	Высокий, ручная постобработка доступна	Можно редактировать в QGIS	Ограничен интерфейсом
Скорость обработки	Быстрая (на локальном ПК)	Зависит от интернета и очереди обработки	Высокая (быстро работает с небольшими файлами)
Подходит для геологических карт	Частично (лучше с чёткими линиями и однородной заливкой)	Да, хорошо работает с цветными картами и сложными формами	Частично, нужна чистка карты перед загрузкой

Приложение Ж

Таблица Ж.1 – Easy tracer характеристики

Раздел	Описание
Совместимость с ГИС	ArcGIS, ArcView, AutoCAD, ГИС «Карта 2003», Credo, MapInfo, MicroStation, WinGIS и другими.
Функциональные возможности	- Сканирование - Геометрическая коррекция - Мощные средства выделения тематических данных - Масочная фильтрация - Сшивка растров - Сложение и вычитание растров и вектора - Автоматическая / полуавтоматическая / ручная векторизация - Ввод атрибутивных данных - Распознавание объектов - Отображение объектов по атрибутивным данным - Автоматическая и полуавтоматическая простановка и контроль высот изолиний - Редактирование векторных примитивов - Генерация и трансформация объектов- Клонирование проектов - Быстрая сборка проектов в общее покрытие - Преобразование типов данных - Преобразование координат по контрольным точкам - Оперативная печать растрово-векторной информации
Входная растровая информация	- Растровые файлы произвольной глубины цветности в форматах: BMP, PCX, CIT, JPG, CALS, TIFF, RLE, DIB
Экспорт / Импорт данных	Векторные форматы: - JET (Easy Trace) - CSV (Comma Separated Values) - DGN (Intergraph) - DXF (AutoCAD) - GEN (ArcInfo) - MIF / MID (MapInfo) - SHP (ArcView) - TOP (Credo) - DWG (AutoCAD) - MAP (Панорама) - MXD; MDB (персональная база геоданных ArcGIS)
Форматы регистрации растров	- TFW - CPT - TAB - MAP

Приложение И

Таблица И.1 – Сравнение по типам ошибки

Тип ошибки	Ручная оцифровка	Полуавтоматическая	Автоматическая
Геометрическая	Небольшие позиционные смещения из-за неточностей оператора, «дрожания руки», неточного масштабирования. Возможно накопление погрешностей при многократной трассировке.	Ошибки геопривязки скана; недостаточное разрешение растрового изображения. Некорректный порог выделения линий приводит к изгибам или смещениям контуров.	Смещение линий из-за шумовых пикселей и ложных цветовых переходов. Алгоритм может «сглаживать» детали, упуская мелкие изломы. “Четность” (parity) ошибки более вероятны при векторизации очень тонких линий.
Топологическая	Разрывы и незамкнутость линий (недоперехлёсты), двойные соединения (перехлёсты), петли и «узлы» из-за случайных поворотов курсора.	Частичное обрыв линий в местах слабого контраста; нестыковки границ смежных полигонов. Автоматические штриховки (например, вода по границам) создают мелкие артефакты. Зачастую требуется последующая доработка топологии вручную.	Алгоритм может вести себя непредсказуемо: создавать «петли» или замыкающие контуры, отсутствующие на карте, как это наблюдалось при использовании ИИ-плагина (формируются «странные» или избыточные полигоны).
Атрибутивная	Ошибки при ручном вводе подписей и классификаций (ошибки транскрипции, опечатки, пропуски меток).	Частично автоматизированное копирование атрибутов (использование шаблонов). Возможны сбои при распознавании текстовых меток (OCR) – пропуски или неправильные символы.	Автоматическая категоризация материалов и площадей может давать ложные совпадения при недостаточном обучении. Цветовые переходы распознаются неточно, из-за чего тип породы определяется неверно.
Программная (ПО)	Практически не характерны (если софт стабилен). Основная роль – удобство редактора (возможность восстановления после ошибки).	Возможны сбои при работе с большими файлами: долгие задержки обработки, зависания ArcScan при сложной топологии. Иногда инструменты векторизации «забывают» автоматически соединить разрывы без дополнительного редактирования.	Высокая зависимость от внешних сервисов (сеть, API). Прерывания связи с сервером ИИ или ошибки формата входа приводят к полному прекращению работы. Ошибочные преобразования могут происходить из-за багов в алгоритме.

РЕЦЕНЗИЯ

на дипломную работу

Мрисиديков Шахнур Хожакбарович

6B07304 – Геопространственная цифровая инженерия

На тему: Автоматизация процессов оцифровки и обработки геопространственных данных с использованием современных ГИС-технологий

Выполнено:

- а) графическая часть на 10 листах
- б) пояснительная записка на 58 страницах

ЗАМЕЧАНИЯ К РАБОТЕ

Рассматриваемая в дипломной работе тема автоматизации процессов оцифровки геопространственных данных является актуальной в условиях цифровой трансформации кадастровой и геоинформационной отраслей. В работе обоснована необходимость перехода от ручной обработки к использованию автоматизированных методов, описаны современные ГИС-платформы (ArcGIS, QGIS), а также сравнены инструменты автоматической векторизации, включая технологии на базе искусственного интеллекта.

Следует отметить следующие положительные стороны дипломной работы: высокий уровень структурированности материала, наличие практической части, связанной с анализом конкретных программных решений, а также попытка интеграции результатов в кадастровую деятельность.

К несущественным замечаниям можно отнести: недостаточно глубокий обзор зарубежных исследований по теме, а также ограниченное количество собственных иллюстративных примеров. Некоторые главы перегружены теоретическими сведениями, при этом разделы с практическими результатами могли бы быть более развёрнутыми.

Оценка работы

В целом дипломная работа представляет собой завершённый исследовательский труд, выполненный в соответствии с требованиями ГОСО. Практическая значимость, актуальность темы и проработка ГИС-инструментария заслуживают положительной оценки.

В связи с этим, дипломная работа заслуживает оценки «отлично» (95%), а студент Мирсидиков Ш.Х., в случае успешной защиты, — присвоения степени бакалавра по специальности «Геопространственная цифровая инженерия».

Рецензент
К.т.н., асс. Профессор
Омиржанова Ж.Т.

« 18 » 26 2025 г.

Подпись Омиржановой Ж.Т.
заварю
HR департамент

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН

Некоммерческое акционерное общество «Казахский национальный исследовательский
технический университет имени К. И. Сатпаева»

Отзыв

научного руководителя на дипломную работу Мирсидикова Ш. Х.

На тему: «Автоматизация процессов оцифровки и обработки геопространственных данных с
использованием современных ГИС-технологий»

Дипломная работа посвящена исследованию методов автоматизированной обработки геопространственных данных на основе современных геоинформационных систем. Актуальность темы определяется необходимостью повышения эффективности пространственного анализа, цифрового картографирования и интеграции данных в кадастровую систему.

Работа носит прикладной характер и ориентирована на решение практических задач по цифровизации геоданных и включению геологических карт в процессы кадастрового учета и территориального планирования.

В процессе подготовки дипломной работы студент продемонстрировал хорошие знания в области ГИС-технологий, цифровой векторизации, обработки пространственной информации и использования программных продуктов QGIS, ArcGIS, Easy tracer и нейросетевых инструментов.

Выпускная работа состоит из двух содержательных глав. В первой главе изложены теоретические основы автоматизации оцифровки геоданных, приведен обзор существующих программных решений и проанализированы их преимущества и ограничения. Во второй главе выполнена практическая часть: студент реализовал сравнение алгоритмов автоматической векторизации, проанализировал точность, производительность и применимость различных ГИС-инструментов.

Работа отличается хорошей структурой, грамотной аргументацией, глубоким анализом и высоким уровнем самостоятельности. Студент показал аналитическое мышление, инициативность и способность применять теоретические знания на практике. Дипломная работа Мирсидикова Ш. Х. заслуживает оценки «отлично», оценивается в 95 баллов и рекомендуется к защите. Студент достоин присуждения степени бакалавра по образовательной программе 6B07304 – Геопространственная цифровая инженерия.

Научный руководитель
PhD, ст. преподаватель



Камза А.Т.

Протокол

о проверке на наличие неавторизованных заимствований (плагиата)

Автор: Мирсидиков Шахнур Хожакбарович

Соавтор (если имеется):

Тип работы: Дипломная работа

Название работы: Автоматизация процессов оцифровки и обработки геопространственных данных с использованием современных ГИС-технологий

Научный руководитель: Анжелика Камза

Коэффициент Подобия 1: 4.2

Коэффициент Подобия 2: 3.1

Микропробелы: 0

Знаки из здругих алфавитов: 2

Интервалы: 0

Белые Знаки: 0

После проверки Отчета Подобия было сделано следующее заключение:

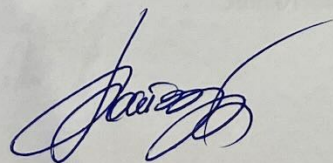
☒ Заимствования, выявленные в работе, является законным и не является плагиатом. Уровень подобия не превышает допустимого предела. Таким образом работа независима и принимается.

☐ Заимствование не является плагиатом, но превышено пороговое значение уровня подобия. Таким образом работа возвращается на доработку.

☐ Выявлены заимствования и плагиат или преднамеренные текстовые искажения (манипуляции), как предполагаемые попытки укрытия плагиата, которые делают работу противоречащей требованиям приложения 5 приказа 595 МОН РК, закону об авторских и смежных правах РК, а также кодексу этики и процедурам. Таким образом работа не принимается.

☐ Обоснование:

Дата 10.06.2022.



Важурбай О.
проверяющий эксперт

Протокол

о проверке на наличие неавторизованных заимствований (плагиата)

Автор: Мирсидиков Шахнур Хожакбарович

Соавтор (если имеется):

Тип работы: Дипломная работа

Название работы: Автоматизация процессов оцифровки и обработки геопространственных данных с использованием современных ГИС-технологий

Научный руководитель: Анжелика Камза

Коэффициент Подобия 1: 4.2

Коэффициент Подобия 2: 3.1

Микропробелы: 0

Знаки из других алфавитов: 2

Интервалы: 0

Белые Знаки: 0

После проверки Отчета Подобия было сделано следующее заключение:

☒ Заимствования, выявленные в работе, является законным и не является плагиатом. Уровень подобия не превышает допустимого предела. Таким образом работа независима и принимается.

☐ Заимствование не является плагиатом, но превышено пороговое значение уровня подобия. Таким образом работа возвращается на доработку.

☐ Выявлены заимствования и плагиат или преднамеренные текстовые искажения (манипуляции), как предполагаемые попытки укрытия плагиата, которые делают работу противоречащей требованиям приложения 5 приказа 595 МОН РК, закону об авторских и смежных правах РК, а также кодексу этики и процедурам. Таким образом работа не принимается.

☐ Обоснование:

Дата 10.06.25г.



Заведующий кафедрой