

АННОТАЦИЯ

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Наименование темы: Разработка сервиса для генерации персональных мелодий на основе анализа эмоционального содержания изображений

Выполнена студентом Базаровым Булатом Баторовичем

Факультет информационных технологий, Новосибирский государственный университет

Кафедра общей информатики

Группа 21201

Направление подготовки 09.03.01 Информатика и вычислительная техника

Направленность (профиль): Программная инженерия и компьютерные науки

Объем работы: 53 страницы

Количество иллюстраций: 9

Количество таблиц: 1

Количество литературных источников: 14

Количество приложений: 2

Ключевые слова: генеративный искусственный интеллект, эмоциональный анализ изображений, мультимодальная интеграция, EmoSet, CLIP, SunoAI

В работе представлен мультимодальный веб-сервис, автоматически генерирующий музыкальные композиции по любым изображениям на основе их эмоционального содержания. Объект исследования — преобразование визуально-эмоциональных признаков в аудиопараметры. Актуальность связана с ростом спроса на персонализированный мультимедийный контент и возможностями генеративных ИИ в объединении визуальных и аудиомодальностей.

Цель работы — разработать сервис, который по загруженному изображению формирует «эмо-вектор» и через облачный API Suno автоматически генерирует соответствующую музыку. Для этого были решены задачи:

1. Обзор существующих решений по генерации музыки и анализу эмоций на изображениях.
2. Обоснование выбора EmoSet, CLIP и Suno API для распознавания восьми базовых эмоций и извлечения перцептивных признаков.
3. Проектирование архитектуры с модулями предобработки изображений, анализа эмоций, генерации текстового промпта, интеграции с API и хранения данных.
4. Реализация ключевых компонентов на Python (FastAPI) и JavaScript (React.js) с обменом по REST и webhook.
5. Разработка методики функционального, нагрузочного и пользовательского тестирования, оценка точности классификации и релевантности музыки.

Создан прототип сервиса, генерирующий в реальном времени мелодии, отражающие эмоциональный фон изображений. Научная новизна — интеграция модели CLIP, дообученной на EmoSet, с синтезом музыки через Suno API и использование расширенного «эмо-вектора» (дискретные эмоции и количественные признаки). Практическая значимость — применение в цифровом творчестве, образовательных и терапевтических приложениях.

Базаров Булат Баторович / _____

20 мая 2025 г.