

Разработка программы распознавания текста с фото при помощи библиотеки Emgu.CV

Батенков Никита Дмитриевич

Приамурский государственный университет им. Шолом-Алейхема

Студент

Аннотация

Целью работы с библиотекой Emgu.CV является разработка приложения по распознаванию текста на фото. Применяются научные материалы, теоретические и эмпирические методы исследования. Практическая значимость работы заключается в создании приложения способного распознавать и выписывать текст с картинки.

Ключевые слова: информационные технологии, искусственный интеллект, распознавание текста, искусственные нейронные сети, машинное обучение

Development of a photo text recognition program using the library Emgu.CV

Batenkov Nikita Dmitrievich

Sholom-Aleichem Priamursky State University

Student

Abstract

The purpose of working with the Emgu library.CV is the development of an application for recognizing text in a photo. Scientific materials, theoretical and empirical research methods are used. The practical significance of the work is to create an application capable of recognizing and writing out text from a picture.

Keywords: information technology, artificial intelligence, text recognition, artificial neural networks, machine learning

1 Введение

1.1 Актуальность

В современном мире текст в электронном виде использовать гораздо удобнее, чем на бумажном носителе, однако не всегда есть возможность получить необходимый текст в нужном формате. В таких случаях могут помочь различные конвейеры файлов.

1.2 Обзор исследований

С.Б.Гусейнов, В.В.Менгель и Д.С.Орозкожоев изучили использование искусственных нейронных сетей в системах, предназначенных для распознавания текста [1].

С.Д.Некрасова описала эксперимент по применению автоматизации на этапе подготовки текста к переводу, с использованием ряда текстов различного формата и размера [2].

Н.Т.Йегемерди, А.Акшабаев рассмотрели технологии распознавания текста, а также использовали методы для повышения эффективности распознавания казахского языка [3].

А.Н.Ганин, М.Н.Голубев разработали программу для ЭВМ, предназначенную для извлечения заданных данных из документов. Программа обрабатывает изображение документа, либо документ в текстовом формате. Программа направлена на обнаружение текстовых фрагментов заданных классов с использованием алгоритмов искусственного интеллекта (компьютерного зрения и обработки естественного языка) [4].

И.А.Волков, А.Л.Исаев в работе рассматриваются и анализируются основные алгоритмы офлайн-распознавания рукописных текстов. В качестве наиболее перспективного предложен алгоритм распознавания на основе структурно-пятенного эталонного метода [5].

С.О.Кулабухова, Ю.В.Бабуров разработали программу, предназначенную для решения задач оптического распознавания текстов (в т.ч. машинопечатных, рукопечатных и рукописных), таблиц, qr кодов, подписей, печатей. ПО может конвертировать в редактируемые форматы отсканированных документов, PDF-документов и файлов изображений документов, включая цифровые фотографии [6].

1.3 Цель исследования

Цель исследования –разработать приложение, распознающее текст с картинки при помощи библиотеки Emgu.CV.

2 Материалы и методы

Программа разработана на языке программирования C# в среде программирования Visual Studio с использованием библиотеки Emgu.CV.

3 Результаты исследования

Для разработки приложения распознающего текст с картинки необходимо: установить библиотеку Emgu.CV, скачать файлы языковых данных, сконструировать окно приложения и запрограммировать элементы меню приложения.

3.1 Установка библиотеки Emgu.CV

В обозревателе решений нажать правой кнопкой мыши по названию проекта в появившемся меню выбрать пункт установка пакетов NuGet в появившемся окне найти в поиске и установить библиотеку Emgu.CV.runtime.windows и Emgu.CV (рис.1).

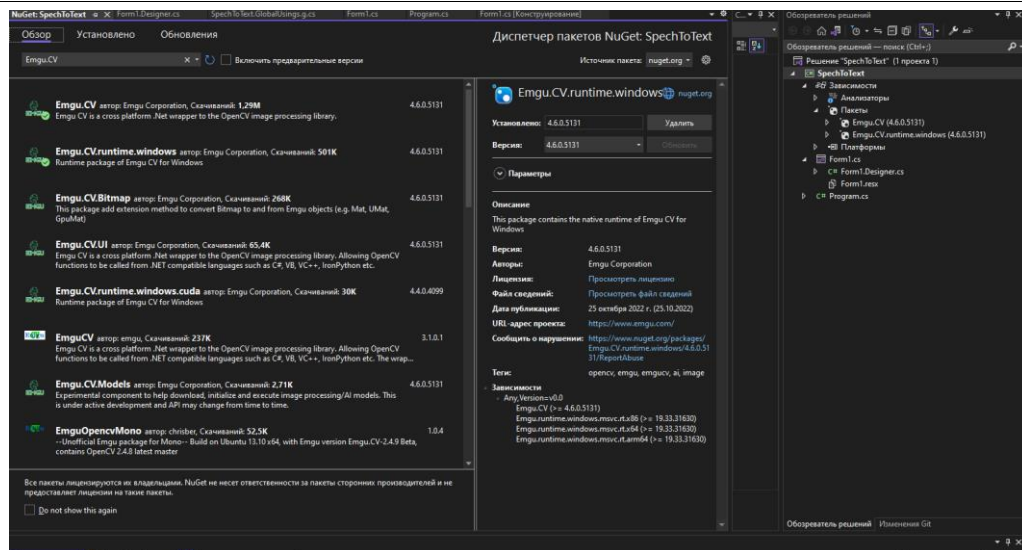


Рисунок 1. Установка библиотеки Emgu.CV

Также для работы программы необходимо скачать файлы языковых данных, работающих с Tesseract 4.0.0

3.2 Конструирование окна приложения

В обозревателе решений нажать правой кнопкой мыши по названию проекта в появившемся меню выбрать пункт добавить, затем создать элемент (рис.2).

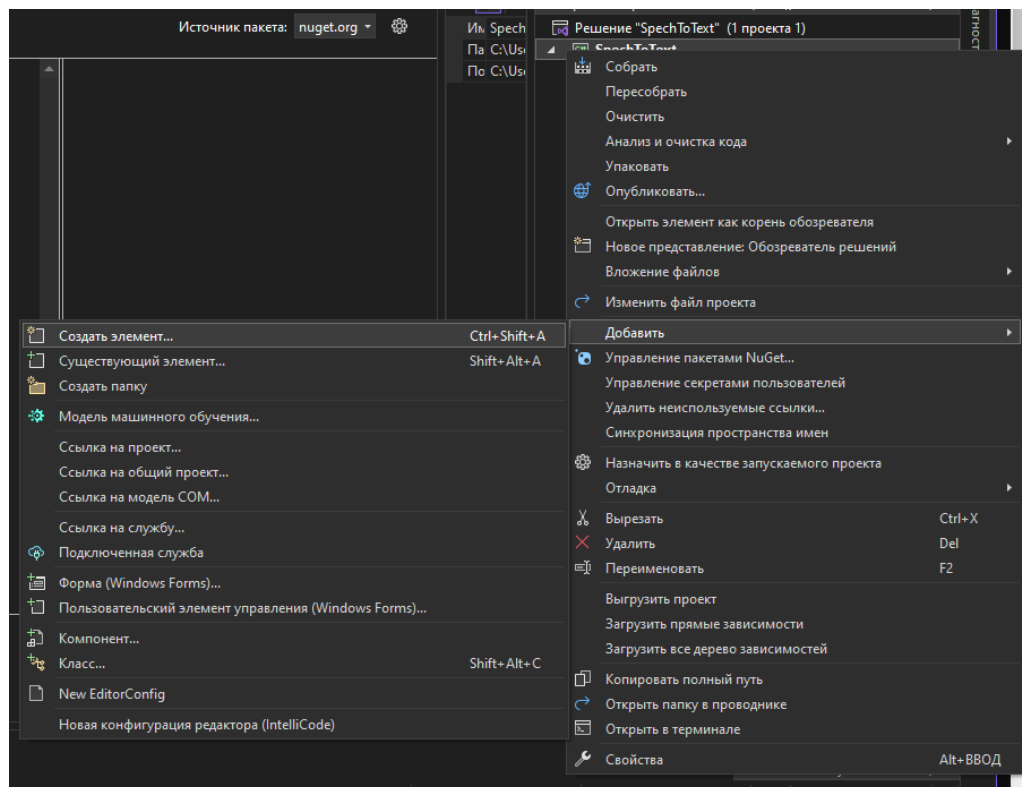


Рисунок 2. Добавления элементов

В открывшемся окне выбрать форму (WindowsForms) после чего создается новая форма на которую необходимо добавить следующие компоненты: menuStrip для размещения кнопки открытия файлов, tools для создания кнопки распознать текст, openFileDialog для взаимодействия с файлами на компьютере, comboBox для выбора языка который будет считан, RichTextBox в который будет записан текст, PictureBox в который будет загружено изображение и TableLayoutPanel на котором будут размещены поля с загруженным фото и полученным текстом (рис.3).

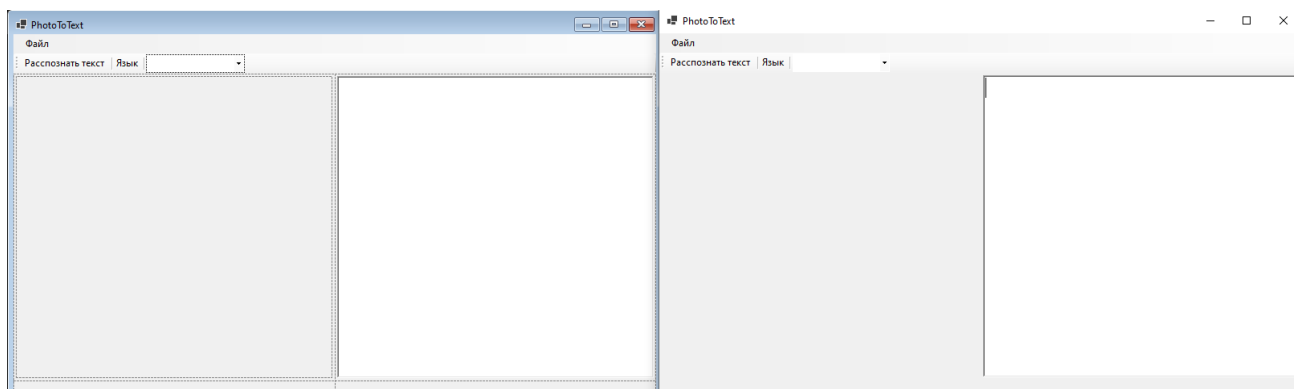


Рисунок 3. Создание формы и результат в действии

3.3 Написание кода

Скрипт выбора картинки с компьютера и вывод выбранной картинки в приложении (рис.4).

```
private void открытьToolStripMenuItem_Click(object sender, EventArgs e)
{
    DialogResult res = openFileDialog1.ShowDialog();
    if (res == DialogResult.OK)
    {
        filePath = openFileDialog1.FileName;
        pictureBox1.Image = Image.FromFile(filePath);
    }
    else
    {
        MessageBox.Show("Картинка не выбрана", "Выберите картинку", MessageBoxButtons.OK, MessageBoxIcon.Exclamation);
    }
}
```

Рисунок 4. Скрипт кнопки открыть

Скрипт выбор языка для распознавания (рис.5).

```
private void toolStripComboBox1_SelectedIndexChanged(object sender, EventArgs e)
{
    if (toolStripComboBox1.SelectedIndex == 0)
    {
        lang = "rus";
    }
    else if (toolStripComboBox1.SelectedIndex == 1)
    {
        lang = "eng";
    }
}
```

Рисунок 5. Скрипт выпадающего списка с выбором языка

Кнопка распознавания текста с загруженной картинкой (рис.6).

```
Ссылка 1
private void toolStripButton1_Click(object sender, EventArgs e)
{
    try
    {
        if (String.IsNullOrEmpty(filePath) || String.IsNullOrWhiteSpace(filePath))
        {
            throw new Exception("картинка не выбрана!");
        }
        else if (toolStripComboBox1.SelectedItem == null)
        {
            throw new Exception("Язык не выбран!");
        }
        else
        {
            Tesseract tesseract = new Tesseract(@"C:\Lang", lang, OcrEngineMode.TesseractLstmCombined);
            tesseract.SetImage(new Image<Bgr, byte>(filePath));
            tesseract.Recognize();
            richTextBox1.Text = tesseract.GetUTF8Text();
            tesseract.Dispose();
        }
    }
    catch (Exception ex)
    {
        MessageBox.Show(ex.Message, "Ошибка", MessageBoxButtons.OK, MessageBoxIcon.Error);
    }
}
```

Рисунок 6. Скрипт обработки фото

3.4 Работа программы

В результате получается следующее приложение (рис.7).

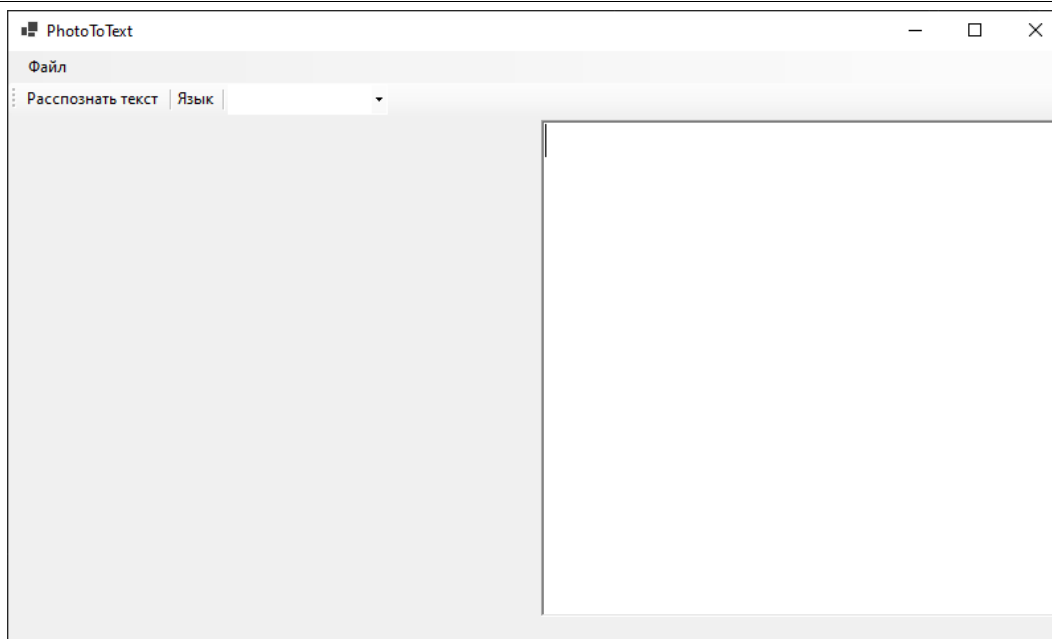


Рисунок 7. Окно готовой программы

При нажатии на элемент меню “Файл” раскрывается меню и можно выбрать “Открыть” для открытия меню выбора файла на компьютере (рис.8).

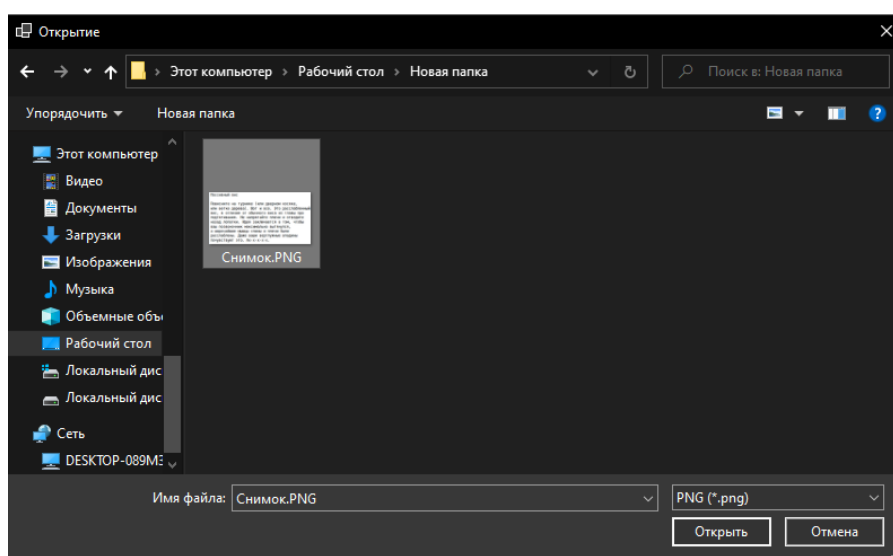


Рисунок 8. Окно “Открытие”

При выборе в окне “Открытие” файла с расширением JPG или PNG, оно загрузится в pictureBox в левой части окна приложения (рис.9).

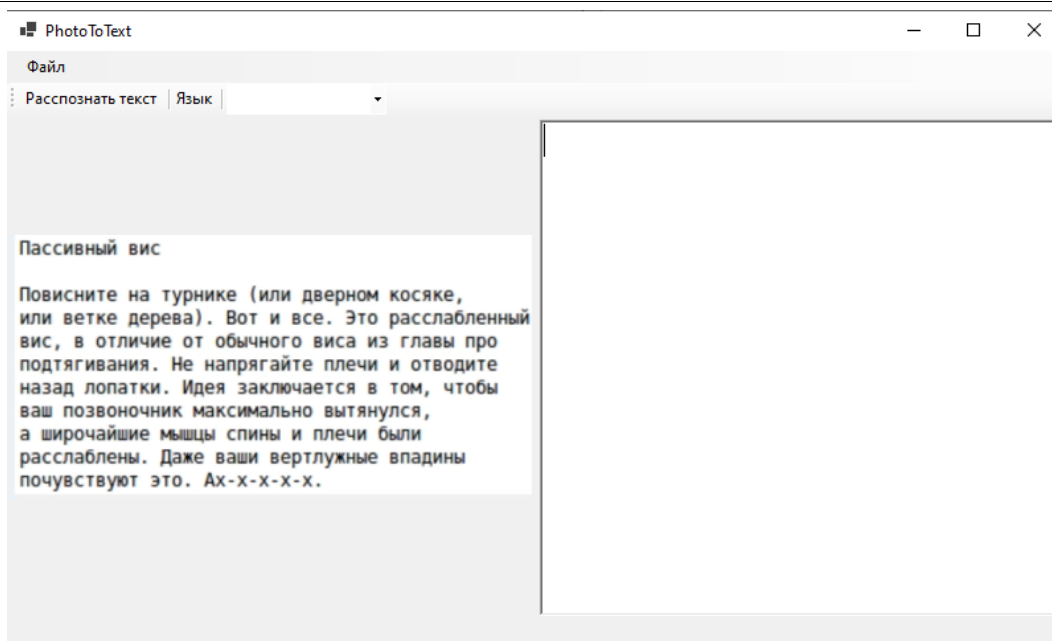


Рисунок 9. Окно приложения с загруженной картинкой

Далее необходимо выбрать язык в выпадающем списке и нажать кнопку “распознать текст” (рис.10). Как можно заметить на рисунке 10, программа считывает текст построчно, что приводит к неправильному переносу текста на новую строку.

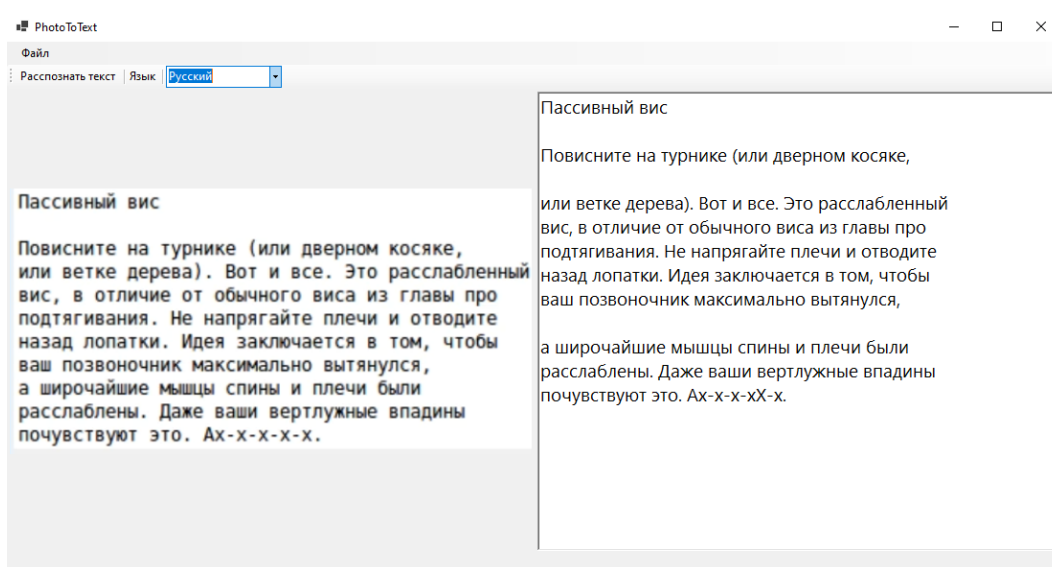


Рисунок 9. Окно приложения с распознанным текстом

Тест программы на более сложном тексте (рис.10), вскрывает множество недостатков построчного считывания текста с картинки, программа не способна отделить текст из разных ячеек превращая результат в набор отрезков предложений. Также видна проблема считывания текстов с символами из нескольких языков из-за которой в тексте появляются неуместные символы.

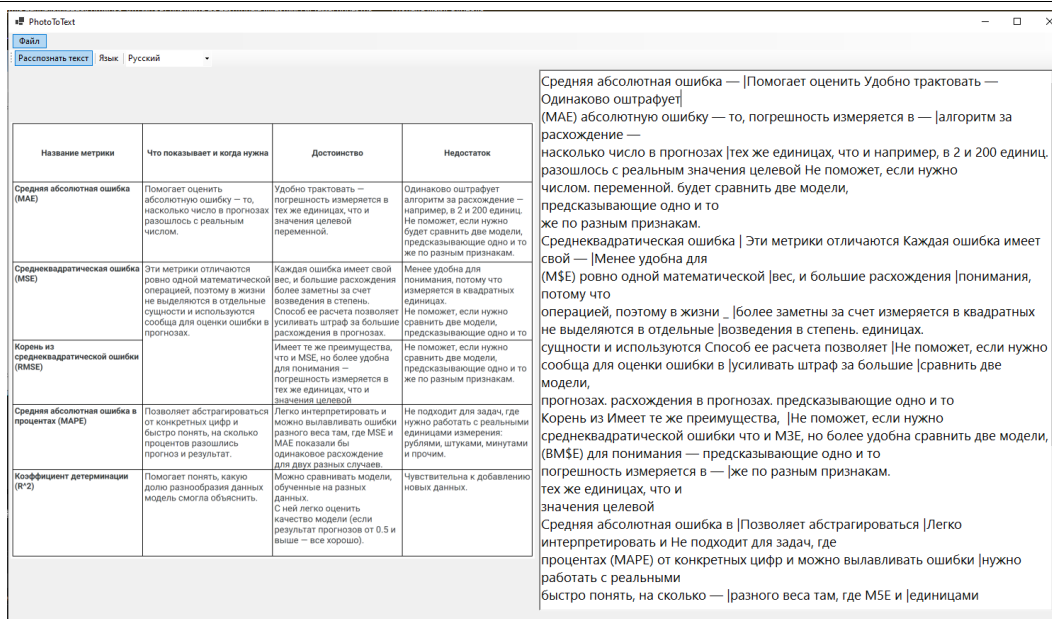


Рисунок 9. Окно приложения с плохо распознанным текстом

4 Вывод

В ходе выполнения работ была разработана программа, распознающая текст с картинки на языке программирования C# при помощи библиотеки Emgu.CV. в среде программирования VisualStudio.

Библиографический список

1. Гусейнов С. Б. Менгель В. В. Орозкожоев Д. С. Использование интеллектуальных систем при распознавании текста на изображении//Известия Кыргызского государственного технического университета им. И. Раззакова.2019. № 1. С. 46 - 52
2. Некрасова С. Д. Технические средства подготовки текста к переводу (на примере программ автоматического распознавания текста) // Проблемы языка и перевода в трудах молодых ученых. 2020. № 19. С. 121-127.
3. Йегемберди Н. Т. Преобразования речи в текст казахского языка для развития поисковых систем // Вопросы устойчивого развития общества. 2022. № 4. С. 1428-1435.
4. Ганин А. Н. Голубев М. Н. Tochka.text - программная платформа для детектирования и распознавания текста на изображениях на основе методов искусственного интеллекта// Свидетельство о регистрации программы для ЭВМ 2021612709, 24.02.2021. Заявка № 2021611995 от 16.02.2021.
5. Волков И. А. Исаев А. Л. Исследование методов распознавания текстов // Modern Science. 2020. № 5-3. С. 552-559.
6. Кулабухова С. О. Бабуров Ю. В. Dreamdocsocr - система оптического распознавания текста Свидетельство о регистрации программы для ЭВМ 2022684779, 16.12.2022. Заявка № 2022684731 от 13.12.2022.