

Решение задачи кластеризации изображений домашних животных с помощью программного пакета визуального программирования Orange

Голубева Евгения Павловна

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

Цель данной статьи – решить задачу кластеризации изображений домашних животных. Для решения задачи кластеризации был использован программный пакет визуального программирования на основе компонентов для визуализации данных Orange и набор данных различных слов. С помощью средств визуализации Orange решили задачу кластеризации изображений домашних животных и получили итоговую схему.

Ключевые слова: Orange, виджет, изображение, кластеризация.

Solving the problem of clustering pet images using the Orange visual programming software package

Golubeva Evgeniya Pavlovna

Sholom-Aleichem Priamursky State University

Student

Abstract

The purpose of this article is to solve the problem of clustering images of pets. To solve the clustering problem, a visual programming software package based on Orange data visualization components and a dataset of various words were used. With the help of Orange visualization tools, we solved the problem of clustering images of pets and obtained the final scheme.

Keywords: Orange, widget, picture, clustering.

1 Введение

1.1 Актуальность

Кластеризация изображений является важным методом анализа данных в области компьютерного зрения. Группировка изображений домашних животных на основе их сходства может предоставить новые знания и привести к пониманию об их или иных особенностях, включая распознавание определенных видов или классов животных.

Программный пакет визуального программирования Orange предоставляет удобный и интуитивно понятный интерфейс для решения задач анализа данных, включая кластеризацию. Благодаря этому пакету исследователи и специалисты в области анализа данных могут легко

применять методы кластеризации к наборам данных, в том числе и к наборам из различных слов.

1.2 Обзор исследований

Н.Ю. Ильясова, А.В. Устинов, А.Г. Храмов проводили экспериментальные результаты кластеризации тестовых и натуральных изображений лейкоцитов крови [1]. Рассматривали этапы решения задач кластеризации изображений с использованием предобученных нейронных сетей А.С. Кузнецов, Е.Ю. Семенов, Л.Д. Матросова [2]. А.С. Титов предложил методику диагностирования с применением алгоритмов кластеризации спутниковых снимков [3]. В статье рассмотрел использование методов кластеризации в программе Orange на основе реальной базы данных. Н. Юсупов [4]. А. В. Леонов в статье рассматривал основные алгоритмы кластеризации категориальных данных применительно к различным типам пользовательских интерфейсов, определяются их достоинства и недостатки [5].

1.3 Цель исследования

Цель исследования - решить задачу кластеризации изображений домашних животных.

2 Материалы и методы

Для решения задачи кластеризации используется программа Orange. Работа будет происходить на готовом наборе данных изображений домашних животных, скачать которые можно по ссылке:

https://drive.google.com/file/d/1BGmkfuJm4xUIhtTbLNq4qRQiQ9O8J_m9/view?usp=sharing

3 Результаты и обсуждения

Перед началом работы требуется установить Orange с официального сайта и установить.

Создадим новый файл (см.рис.1).



Рисунок 1 - Создание нового файла

Для решения задачи класстеризации изображения необходимо установить дополнение Image Analytics. Для того, чтобы скачать дополнение, необходимо перейти в Options, далее в Add-ons, в появившемся окне выбираем Image Analytics (см.рис.2).

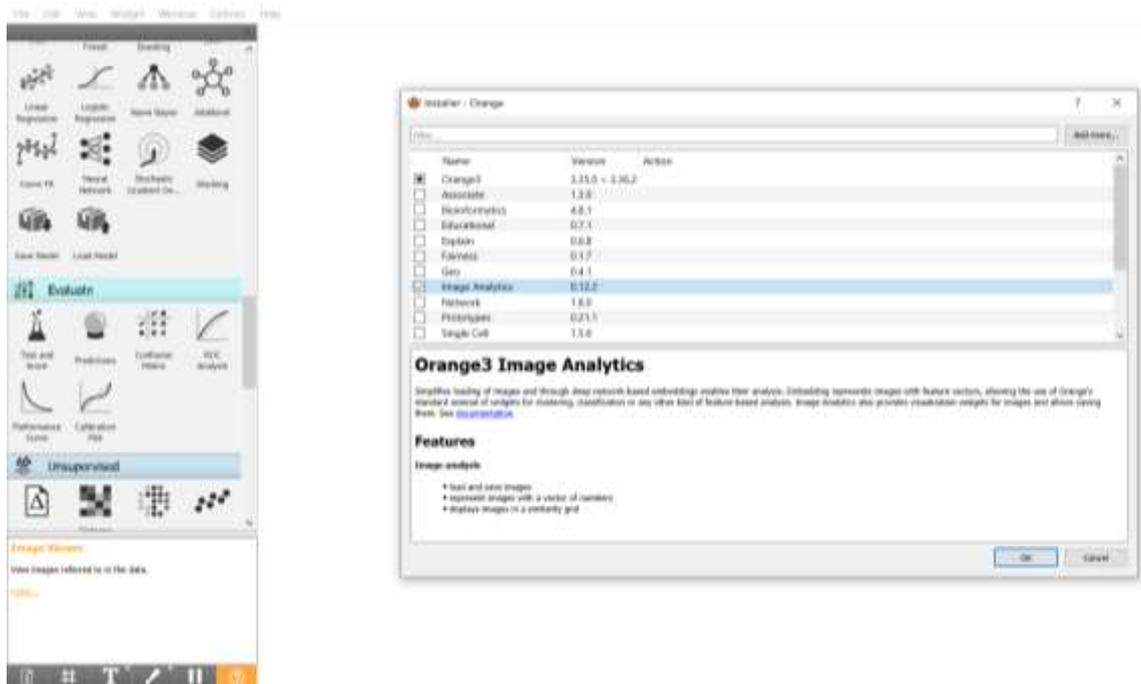


Рисунок 2 - Установка дополнения Image Analytics

Для того, чтобы загрузить изображения, необходимо из раздела Image Analytics выбрать виджет Import Images и перенести его на холст (см.рис.3).



Рисунок 3 - Добавление виджета Import Images на холст

Открываем виджет Import Images и подгружаем подготовленную папку с изображениями домашних животных. После того как подгрузили папку появилась информация, что папка с изображениями животных содержит и 50 картинок (см.рис.4).



Рисунок 4 - Загрузка папки с изображениями домашних животных

Для того чтобы посмотреть загруженные изображения добавляем виджет Image Viewer, и соединяем с виджетом Import Images (см.рис.5).

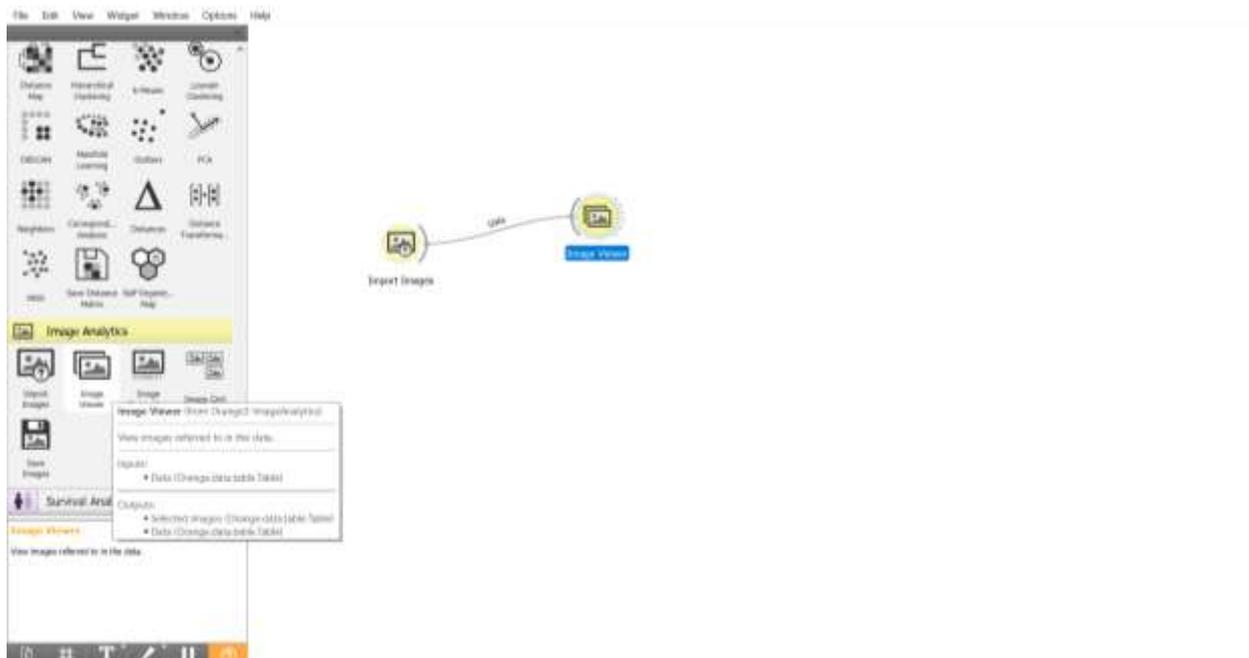


Рисунок 5 - Добавление виджета Image Viewer

Открываем виджет Image Viewer. В появившемся окне можно увидеть изображение животных (см.рис.6).



Рисунок 6 - Просмотр изображений

Далее добавим виджет Image Embedding на холст и соединим с виджетом Import Images (см.рис.7). Виджет Image Embedding получает таблицу с изображениями и отправляет их на сервер для встраивания в формате, понятном алгоритмам машинного обучения. Сервер проталкивает изображения через предварительно обученную глубокую нейронную сеть и возвращает числовые векторы в виджет.

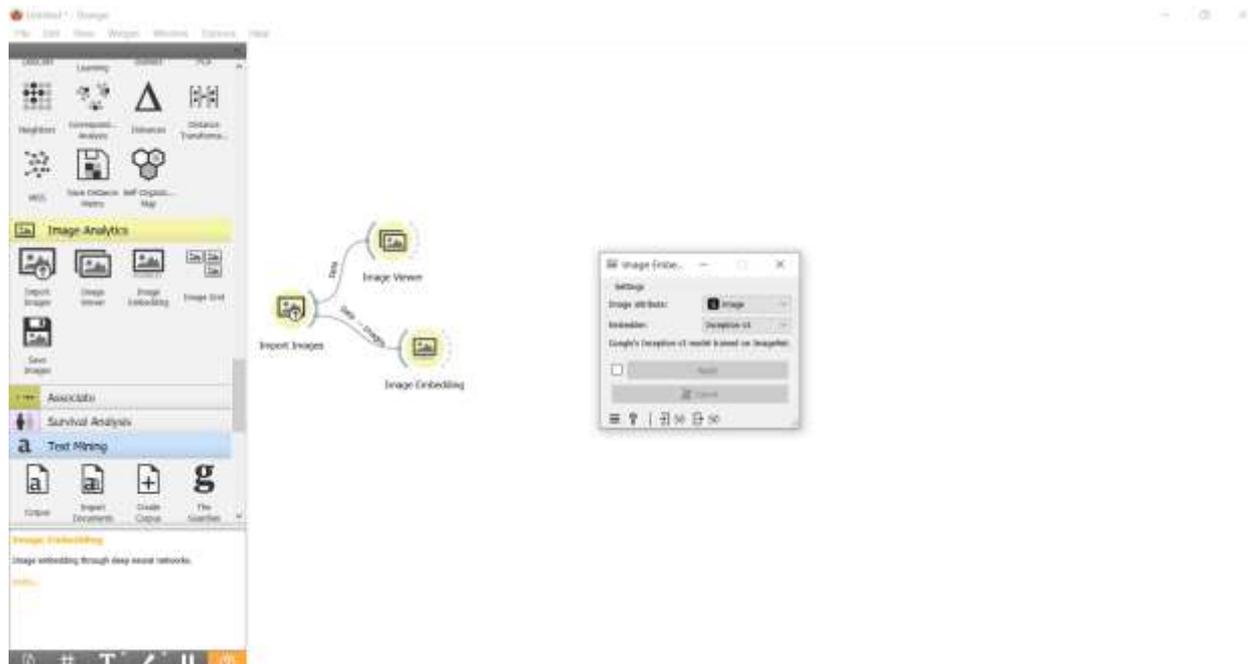


Рисунок 7 - Добавление виджета Image Embedding

Для того чтобы просмотреть данные виджета Image Embedding, добавляем виджет Data Table и соединяем с Image Embedding (см.рис.8).

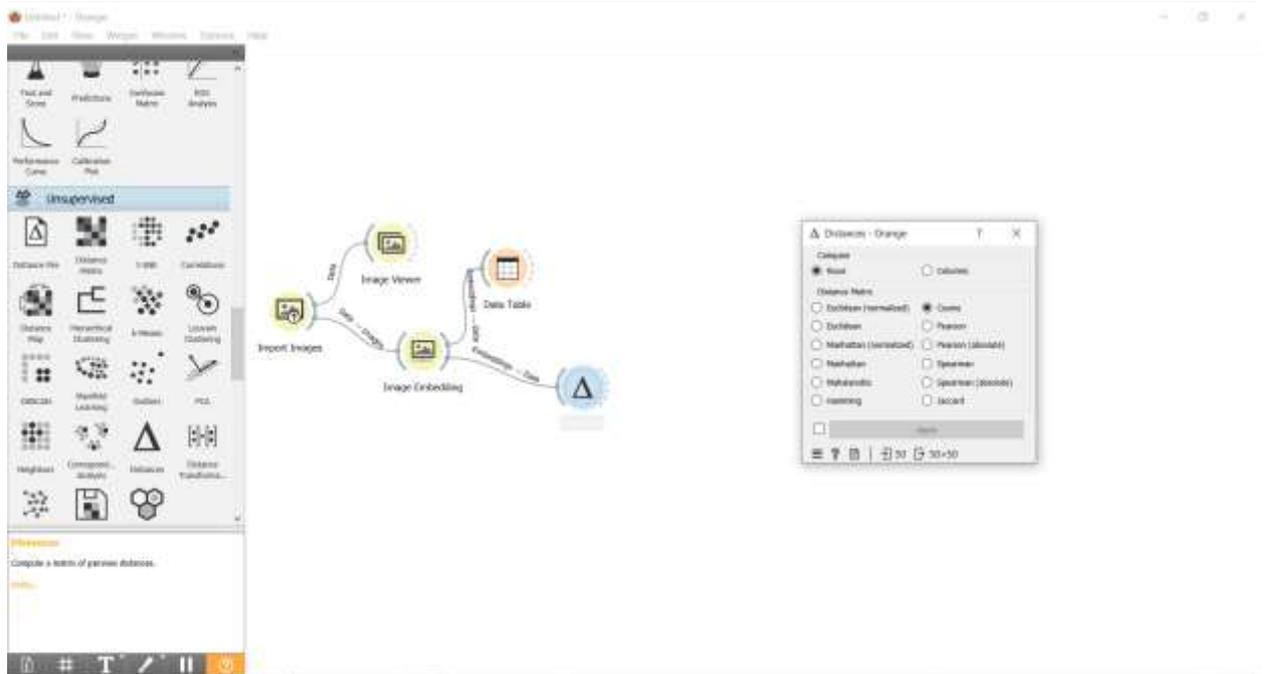


Рисунок 10 - Добавление виджета Distances на холст

Далее проведем кластеризацию изображений. Для этого добавим виджет Hierarchical Clustering на холст и соединим с виджетом Distances (см.рис.11).

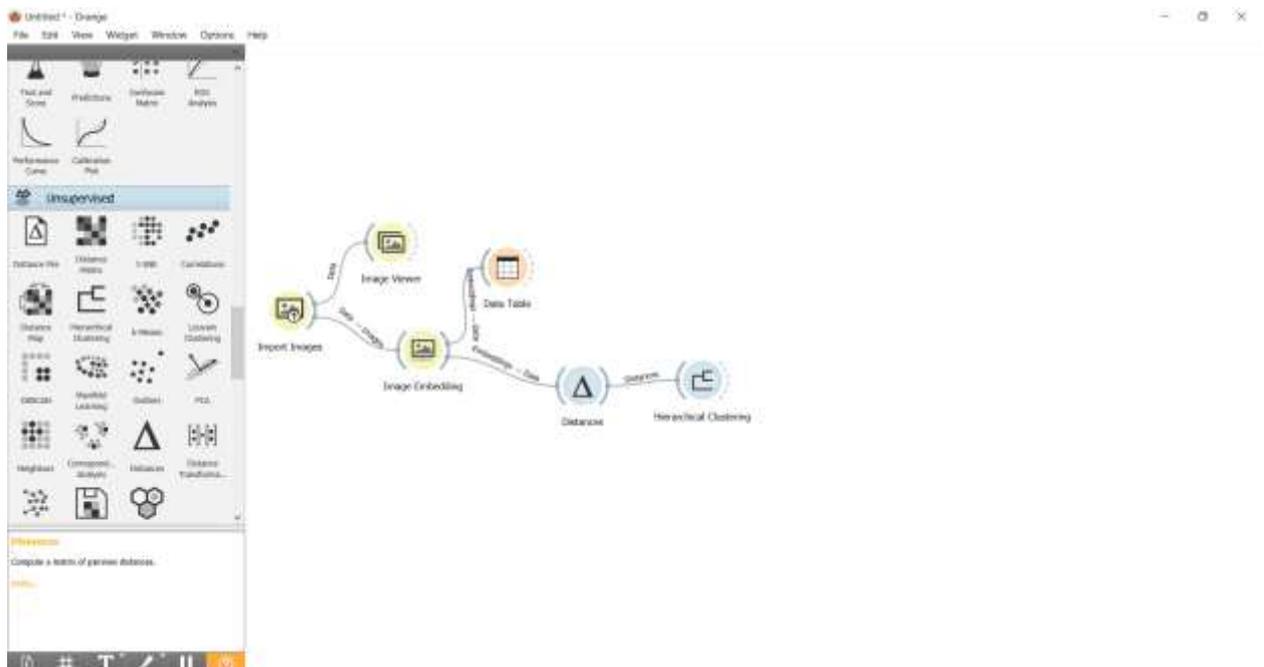


Рисунок-11 Добавление виджета Hierarchical Clustering на холст

Открываем виджет Hierarchical Clustering. В появившемся окне можно увидеть, как изображения распределились на 5 кластеров (см.рис.12).

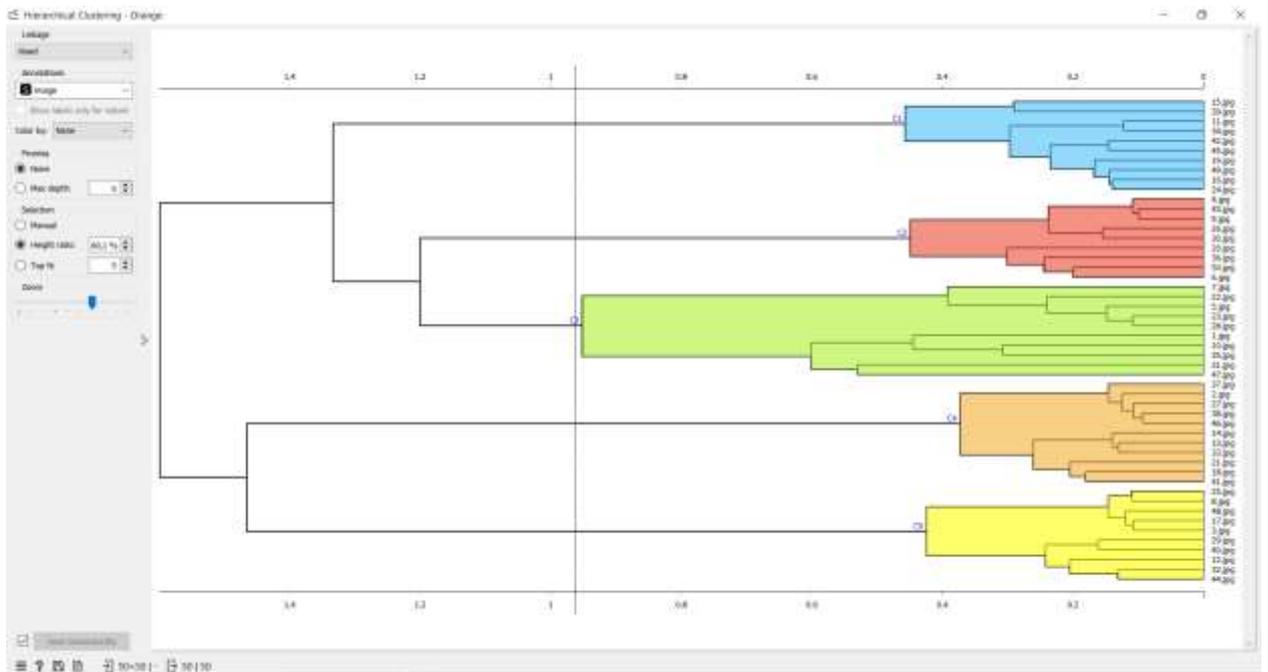


Рисунок 12 - Распределение изображений по кластерам

Для того чтобы убедиться, что изображения распределены правильно по кластерам, добавим виджет Image Viewer и соединяем с виджетом Hierarchical Clustering (см.рис.13).

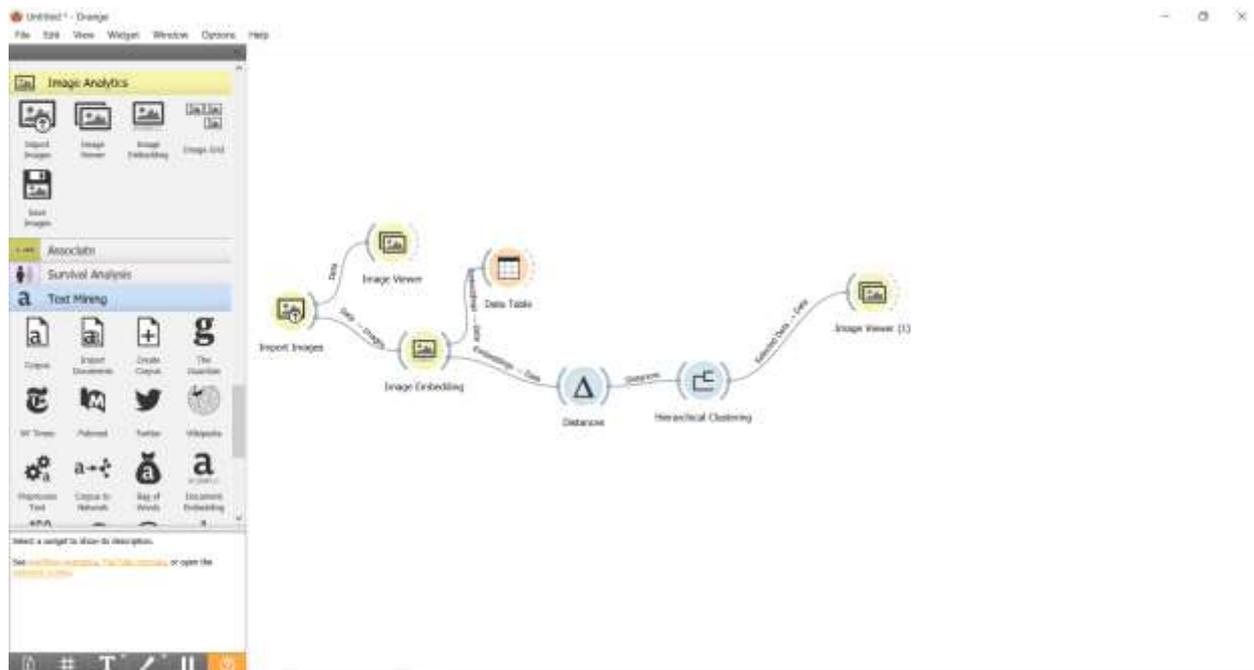


Рисунок 13 - Добавление виджета Image Viewer на холст

Для того чтобы убедиться в правильности результата, откроем виджет Image Viewer. В Hierarchical Clustering выберем результат c1, и изображения автоматически появятся в Image Viewer. В окне Image Viewer мы наблюдаем, что действительно результат c1 содержит изображения кошек (см.рис.14).

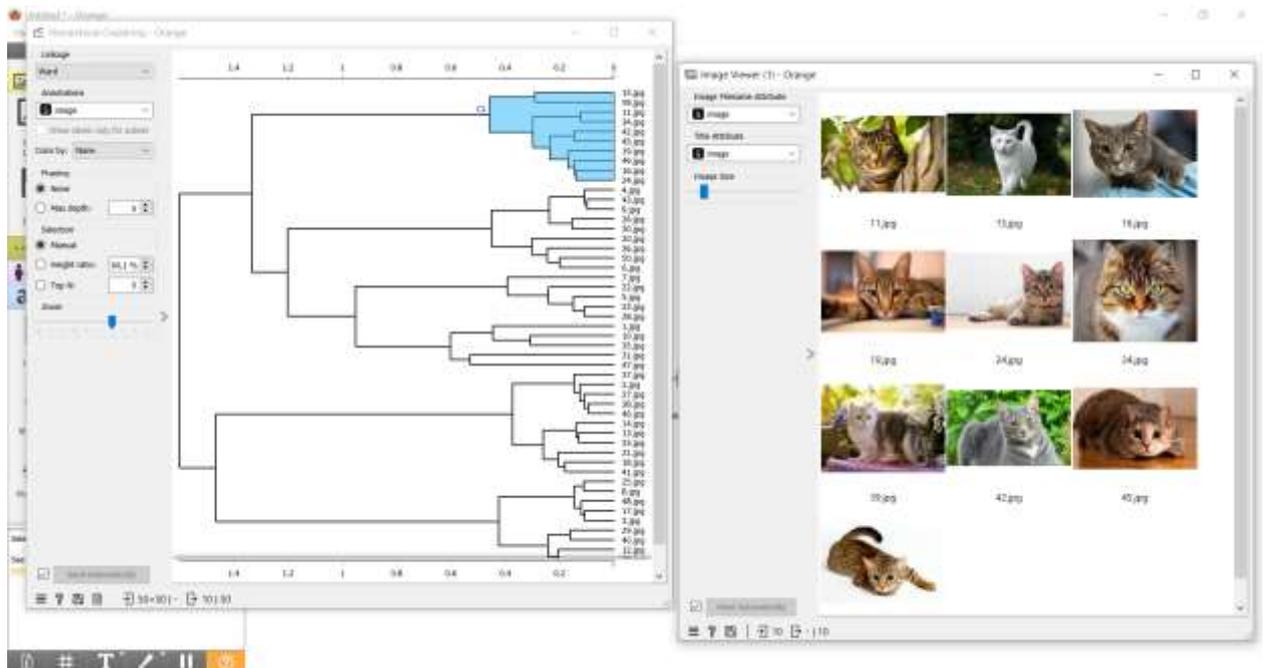


Рисунок 14 - Проверка результатов

В итоге получилась готовая схема, с помощью которой можно решить задачу кластеризации изображений (см.рис.15).

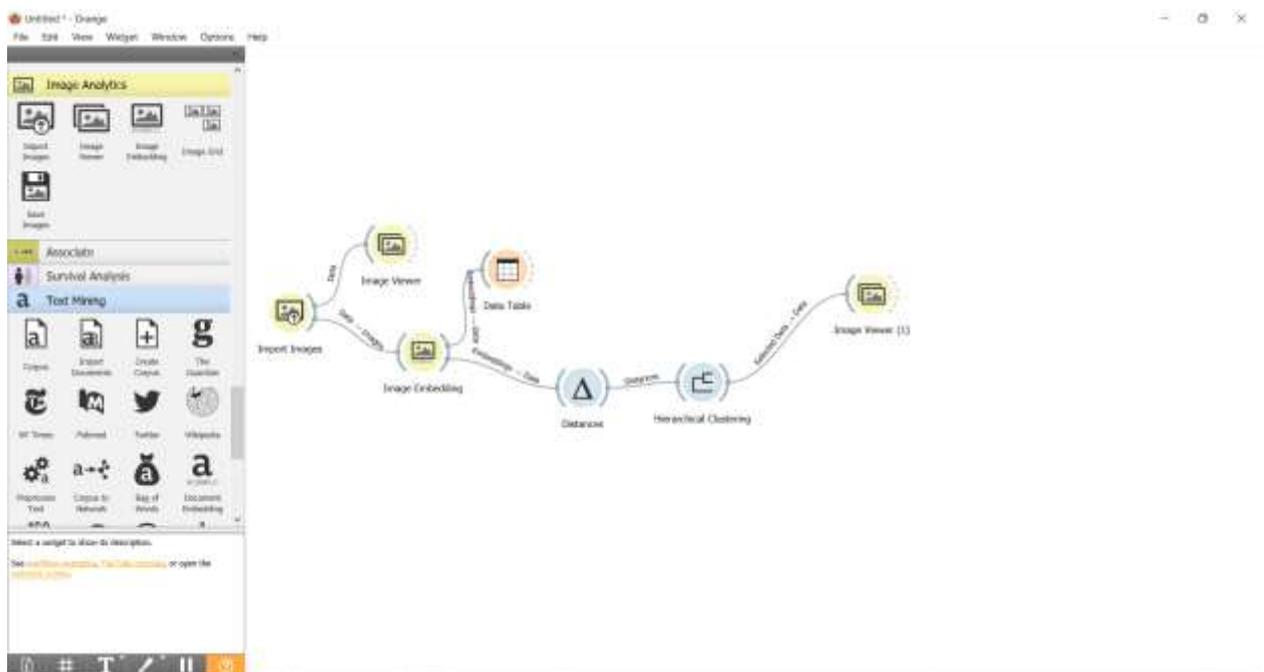


Рисунок 15 - Итоговая схема

Выводы

В данной работе была выполнена задача кластеризации набора данных изображений домашних животных с помощью программного пакета визуального программирования на основе компонентов для визуализации данных Orange. С помощью виджетов Import Image, Image Viewer, Image Embedding, Data Table, Distances, Hierarchical Clustering выполнили

кластеризацию изображений домашних животных и получили итоговую схему.

Библиографический список

1. Кузнецов А. С., Семенов Е. Ю., Матросова Л. Д. Кластеризация изображений при использовании предобученных нейронных сетей //International journal of open information technologies. 2019. Т. 7. №. 7. С. 42-47.
2. Моисеенко Г. А. и др. Классификация и распознавание изображений живой и неживой природы //Оптический журнал. 2015. Т. 82. №. 10. С. 53-64.
3. Юсупов Н. Исследование методов кластеризации в программе Orange //Молодежная школа-семинар по проблемам управления в технических системах имени АА Вавилова. 2020. Т. 1. С. 35-37.
4. Клименко А. В., Слащев И. С. кластерный анализ данных //Вестник науки. 2019. Т. 1. №. 1. С. 159-163.
5. Гринченков Д. В. и др. Сравнительный анализ алгоритмов интеллектуального анализа данных //Моделирование. Теория, методы и средства. 2016. С. 263-266.
6. Мастевой С. С., Петрова А. Н. Data mining: обзор методов и области их применения // Наука, инновации и технологии: от идей к внедрению. 2022. С. 38-40.
7. Ильясова Н. Ю. и др. Алгоритмы автоматической кластеризации изображений микрочастиц //Компьютерная оптика. 1993. №. 13. С. 39-46.