Приложение 1

к договору № \_\_\_\_

|  |  |
| --- | --- |
| **СОГЛАСОВАНО**  \_\_\_\_\_\_\_\_\_\_\_\_\_ С.М. Липкин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2022 г. | **УТВЕРЖДАЮ**  \_\_\_\_\_\_\_\_\_\_\_\_\_ Н.В. Старостин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2022 г. |

**ПО «Micro atomic»**

**Руководство программиста**

**№ 02068143.00237-04**

Оглавление

[Введение 3](#_Toc107187108)

[1 Назначение и условия применения программы 3](#_Toc107187109)

[1.1 Структура ПО 3](#_Toc107187110)

[1.2 Подсистема «DataParser» 3](#_Toc107187111)

[1.3 Подсистема «ResultsPrinter» 3](#_Toc107187112)

[1.4 Подсистема «OutputHandler» 4](#_Toc107187113)

[1.5 Библиотека «Neural Network» 4](#_Toc107187114)

[1.6 Фреймворк «Keras 2.4.0 (TensorFlow 2.4.1)» 5](#_Toc107187115)

[1.7 Фреймворк «NumPy 1.19.5» 5](#_Toc107187116)

[2 Характеристики программы 5](#_Toc107187117)

[2.1 Характеристики подсистемы «DataParser» 5](#_Toc107187118)

[2.2 Характеристики подсистемы «ResultsPrinter» 5](#_Toc107187119)

[2.3 Характеристики подсистемы «OutputHandler» 5](#_Toc107187120)

[2.4 Характеристики библиотеки «Neural Network» 6](#_Toc107187121)

[3 Обращение к программе 6](#_Toc107187122)

[3.1 Обращение к подсистеме « DataParser» 6](#_Toc107187123)

[3.2 Обращение к подсистеме « ResultsPrinter» 6](#_Toc107187124)

[3.3 Обращение к подсистеме « OutputHandler» 6](#_Toc107187125)

[4 Входные и выходные данные подсистем 6](#_Toc107187126)

[4.1 Входные и выходные данные подсистемы «DataParser» 6](#_Toc107187127)

[4.2 Входные и выходные данные подсистемы «ResultsPrinter» 6](#_Toc107187128)

[4.3 Входные и выходные данные подсистемы « OutputHandler» 6](#_Toc107187129)

# Введение

Данное руководство описывает структуру библиотеки, обеспечивающей функционирование системы для определения наличия «зародыша» на поверхности металла «Анализ данных атомной силовой микроскопии», которая обеспечивает обнаружение существования зародыша на снимке атомного силового микроскопа. Приводятся основные принципы работы и взаимодействия с подсистемами консольного приложения.

# Назначение и условия применения программы

## Структура ПО

Основными подсистемами, обеспечивающими функционирование ПО, являются:

1. DataParser
2. ResultsPrinter
3. OutputHandler
4. Neural Network;

Основными библиотеками и фреймворками, обеспечивающими функционирование ПО являются:

1) Keras (TensorFlow 2);

2) NumPy;

## Подсистема «DataParser»

Подсистема “DataParser” предназначена для разбора входных данных в формате NOVA и перевода их во внутреннее представление системы.

Подсистема содержит класс “DataCollector”, содержащий поля, хранящие метаданные переданного на вход изображения в формате NOVA.

“DataCollector” содержит следующие методы:

* Конструктор – ответственен за создание класса “DataCollector”
* ReadFile(path) – считывает данные из файла, путь к которому содержится в параметре path
* ReadImage(path) - считывает данные из картинки, путь к которому содержится в параметре path
* \_\_ParseData() – заполняет поля класса соответствующими данными на основе переданного на вход изображения в формате NOVA.

Главным условием функционирования подсистем является соблюдение формата входных данных. Подсистемы способна считать входные данные в формате NOVA (снимок атомного силового микроскопа).

## Подсистема «ResultsPrinter»

Подсистема “ResultsPrinter” предназначена для отображения информации о полученных результатах в консоли.

Подсистема содержит функции, которые выводят в консоль предсказание нейронной сети, вероятность правильности предсказания, время за которое сеть предсказала результат, а также время считывания входных данных).

Главными условиями применения подсистемы являются соблюдение контрактов программных объектов подсистемы. Подсистема не предъявляет никаких требований к наличию периферийных устройств. Подсистема не предъявляет никаких собственных требований к оперативной памяти или иным вычислительным ресурсам.

## Подсистема «OutputHandler»

Подсистема “OutputHandler” предназначена для записи результатов работы в текстовом файле.

Подсистема содержит функции, которые выводят в файл формата txt при использовании флага запуска «-f» предсказание нейронной сети, вероятность правильности предсказания, время за которое сеть предсказала результат, а также время считывания входных данных).

Главными условиями применения подсистемы являются соблюдение контрактов программных объектов подсистемы. Подсистема не предъявляет никаких требований к наличию периферийных устройств. Подсистема не предъявляет никаких собственных требований к оперативной памяти или иным вычислительным ресурсам.

## Библиотека «Neural Network»

Библиотека содержит подсистемы: “Data Parser”, “Data Converter” и “NN Prototype”

Подсистема “Data Parser” предназначена для разбора входных данных (изображения в формате NOVA) и перевод их во внутреннее представление системы. Эта подсистема определяет объект, предназначенный для хранения «первичных» данных в удобном для последующей работы программы формате.

Под «первичными» данными будем понимать данные, считанные непосредственно из входной информации (возможно, после предварительной обработки), под «вторичными» - данные, вычисляемые либо непосредственно подсистемой на основе «первичных» данных, либо вычисляемые решающей частью и сохраняемые в подсистеме.

Функциональные обязанности подсистемы “Data Parser”:

1. Хранение «первичных» данных о изображении в формате NOVA.

Главным условием функционирования подсистем является соблюдение формата входных данных.

Подсистема “Data Converter” предназначена для обеспечения функционирования решающей части системы. Эта подсистема определяет логику обработки и вычисления «вторичных данных», также запрашиваемых решающей частью.

Функциональные обязанности подсистемы “Data Converter":

1. Нормализация «первичных» данных, преобразование их во «вторичные».

Главными условиями применения подсистемы являются соблюдение контрактов программных объектов подсистемы. Подсистема не предъявляет никаких требований к наличию периферийных устройств, а также к объемам оперативной памяти и прочим характеристикам вычислительной системы, поскольку является «пассивной» программной подсистемой.

Подсистема “NN Prototype” предназначена для создания и тренировки модели нейронной сети. Такая нейронная сеть может определить наличие зародыша в переданном на вход изображении.

Система содержит следующие функции:

* SaveModel(model) – сохраняет модель “model” в файл в формате H5;
* PrintModelStructure(model) – выводит структуру модели “model”;
* CreateModel(image\_x, image\_y, learning\_rate) – создает модель нейронной 'сети, учитывая размер входного слоя [image\_x, image\_y], а также переданный параметр “learning\_rate”.

Функциональные обязанности подсистемы:

1. Сохранение модели нейронной сети в формате H5;
2. Вывод структуры нейронной сети;
3. Чтение данных для обучения нейронной сети (датасета);
4. Тренировка модели нейронной сети;
5. Тестирование модели нейронной сети;

Главными условиями применения подсистемы являются соблюдение контрактов программных объектов подсистемы. Подсистема предъявляет особые требований к объему оперативной памяти, видео памяти и производительности вычислительной системы. Минимальные требования: процессор x86\_64 с тактовой частотой 3 ГГц, оперативная память DDR4 объемом 8 Гб, видеокарта с типом памяти GDDR5, частотой памяти 8000МГц и объемом видеопамяти 6Гб.

## Фреймворк «Keras 2.4.0 (TensorFlow 2.4.1)»

Keras — открытая библиотека, написанная на языке Python и обеспечивающая взаимодействие с искусственными нейронными сетями. Она представляет собой надстройку над фреймворком TensorFlow. Нацелена на оперативную работу с сетями глубинного обучения, при этом спроектирована так, чтобы быть компактной, модульной и расширяемой.

Применяется для упрощения работы с моделью нейронной сети.

## Фреймворк «NumPy 1.19.5»

NumPy — открытая библиотека, написанная на языке Python, обеспечивающая поддержку многомерных массивов (включая матрицы) и поддержку высокоуровневых математических функций, предназначенных для работы с многомерными массивами.

Применяется для упрощения работы с массивами данных.

# Характеристики программы

## Характеристики подсистемы «DataParser»

Главным фактором нормального функционирования подсистемы является соответствие входных данных формату входных данных. В случае несоответствия или в случае нарушения информационной структуры входных данных подсистема также генерирует исключения, которые должны быть обработаны клиентским кодом соответствующим образом.

## Характеристики подсистемы «ResultsPrinter»

Главным фактором нормального функционирования подсистемы является выполнение контрактов подсистемы и наличие соответствующего окружения.

Система определяет систему классов, позволяющих выводить результаты работы программы в консоль.

## Характеристики подсистемы «OutputHandler»

Главным фактором нормального функционирования подсистемы является выполнение контрактов подсистемы и наличие соответствующего окружения.

Система определяет систему классов, позволяющих записывать результаты работы программы в текстовый файл.

## Характеристики библиотеки «Neural Network»

Главными условиями применения подсистемы являются соблюдение контрактов программных объектов подсистемы. Подсистема предъявляет особые требований к объему оперативной памяти, видео памяти и производительности вычислительной системы. Минимальные требования: процессор x86\_64 с тактовой частотой 3 ГГц, оперативная память DDR4 объемом 8 Гб, видеокарта с типом памяти GDDR5, частотой памяти 8000МГц и объемом видеопамяти 6Гб.

# Обращение к программе

## Обращение к подсистеме « DataParser»

Поскольку внешние обязанности подсистемы невелики, то и контракт обращения к системе состоит всего из одного класса – DataCollector. Объект этого класса получает на вход имя файла, содержащего входные данные, а на выходе выдает сконфигурированный соответствующим образом объект, готовый к передаче библиотеке “Neural Network”.

## Обращение к подсистеме « ResultsPrinter»

Подсистема определяет единственный класс ResultsPrinter, обеспечивающий вывод информации о результатах работы нейронной сети, в консоль.

## Обращение к подсистеме « OutputHandler»

Подсистема определяет единственный класс OutputHandler, обеспечивающий вывод полной информации о результатах работы нейронной сети, в текстовый файл.

# Входные и выходные данные подсистем

## Входные и выходные данные подсистемы «DataParser»

Входными данными является изображение в формате NOVA, выходными данными является объект класса NOVA. Parser, описанный в данной подсистеме, содержащий информацию о переданном на вход изображении.

## Входные и выходные данные подсистемы «ResultsPrinter»

Данная подсистема работает в двух режимах:

* Вывод предсказаний;
* Вывод валидации.

В режиме предсказаний входными данными является ответ нейронной сети о наличии зародыша, который представляет из себя скаляр, являющийся вероятностью наличия зародыша.

В режиме вывода валидации подсистема получает на вход информацию о средней точности предсказания нейронной сети на данных, представляющих из себя набор из 500 входных изображений.

## Входные и выходные данные подсистемы « OutputHandler»

Входными данными является ответ нейронной сети о наличии зародыша, который представляет из себя скаляр, являющийся вероятностью наличия зародыша.