

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Комсомольский-на-Амуре государственный университет»

На правах рукописи

Карамышев Максим Дмитриевич

**Проектирование и разработка программного обеспечения
для определения параметров заготовки с помощью
технологий машинного обучения и компьютерного зрения**

Направление подготовки
09.04.01 «Информатика и вычислительная техника»

**АВТОРЕФЕРАТ
МАГИСТЕРСКОЙ ДИССЕРТАЦИИ**

2023

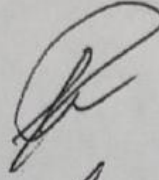
Коротченко
Лариса Никитовна

Проверено

03.07.2023 Зачтено Библиотека

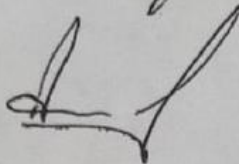
Работа выполнена в ФГБОУ ВО «Комсомольский-на-Амуре
государственный университет»

Научный руководитель:



кандидат технических наук, доцент, зав.
кафедрой ПУРИС ФГБОУ ВО «Комсо-
мольский-на-Амуре государственный
университет» Петрова Анна Николаевна

Рецензент:

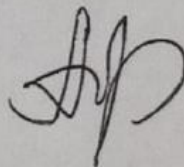


кандидат физико-математических наук,
доцент кафедры ИБИСиФ ФГБОУ ВО
«Амурский гуманитарно-педагогиче-
ский государственный университет»
Анисимов Антон Николаевич

Защита состоится «29» июня 2023 года в 14 часов 00 мин. на заседании
государственной экзаменационной комиссии по направлению 09.04.01 «Информа-
тика и вычислительная техника» в Комсомольском-на-Амуре государственном
техническом университете по адресу: 681013, г. Комсомольск-на-Амуре, пр. Ле-
нина, 27, ауд. 313/5

Автореферат разослан 23 июня 2023 г.

Секретарь ГЭК



Е.В. Абрамсон

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. На данный момент на предприятии «Амурсталь» существует проблема, связанная с процессом измерения заготовок. Проблема состоит в том, что на данный момент процесс измерения заготовок происходит вручную. Проблема затрагивает сотрудников отдела технологического контроля (ОТК), которые занимаются составлением сертификата качества заготовок и сотрудников участка машины непрерывного литья заготовки (МНЛЗ), которые ответственны за нарезание заготовок. Ручное измерение приводит к тому, что сертификат качества может заполняться некорректными данными, что в дальнейшем может привести к проблемам на дальнейших стадиях производства. Удачным решением была бы разработка системы автоматизированного измерения заготовок с помощью алгоритмов компьютерного зрения.

При обработке изображений одной из центральных является задача распознавания заданных объектов. Алгоритмы обработки изображений все чаще используются в научных и прикладных исследованиях в различных областях человеческой деятельности. В связи с этим, задачи анализа изображений - одни из наиболее важных при обработке изображений.

Значительный рост вычислительной мощности компьютеров и появление новых математических моделей и алгоритмов в последние годы позволило добиться значительного прогресса в области компьютерного зрения. Новые технологии, основанные на машинном обучении, позволяют создавать надежные решения для сложных задач распознавания зрительных образов. Многие из этих решений имеют практические применения на рынке, а потому привлекают все больше инвестиций.

Применение компьютерного зрения в различных сферах жизни уже не ново и ему посвящено множество исследований как российских ученых: Ермилов Д. Ю Венцов [1], Н. Н. Гибридный [2], Сергиенко [3], Ледяев В. П [4]; так и зарубежных: D. Forsite, J. Pons [5], Goodfellow, Y. Bengio, A. Courville [6], M. Nielsen [7], W. Dai, C. Dai, S. Qu, J. Li, S. Das [8], J. Yosinski [9], L. Shapiro, J. Stockman [10]. На данный

момент существует большое количество реализаций технологии компьютерного зрения, как универсальных, так и направленных на решение конкретных задач.

Цель диссертационной работы является разработка алгоритмов и их реализация для уменьшения затрат на определение размера заготовки.

Задачи исследования. Для достижения поставленной цели необходимо решить следующие задачи:

- 1) исследовать предметную область;
- 2) изучить современные методы и инструменты обработки и анализа изображения;
- 3) изучить алгоритмы обработки изображений с помощью глубокого машинного обучения и выбрать наиболее эффективные;
- 4) изучить алгоритмы компьютерного зрения и выбрать наиболее подходящие для решения поставленной задачи;
- 5) определить список необходимого функционала для разрабатываемого приложения;
- 6) проанализировать аналоги ПО, имеющего схожий функционал;
- 7) спроектировать модель взаимодействия программного продукта с пользователем на основе сформированного списка функционала;
- 8) проанализировать и выбрать оптимальные средства разработки приложения;
- 9) изучить разработку приложений с применением библиотеки OpenCV;
- 10) разработать приложение в соответствии со спроектированной ранее моделью;
- 11) протестировать разработанное ПО и устранить найденные ошибки.

Объект исследования. Программное обеспечение, предназначенное для измерения объектов с помощью компьютерного зрения.

Предмет исследования. Точность алгоритмов по нахождению и определению длины заготовки, разработанных с помощью технологий компьютерного зрения и глубокого машинного обучения используемые для анализа заготовки.

Методы исследования:

- Теоретические методы: используются методы математического и компьютерного моделирования.
- Эмпирические методы: сравнение различных методов по анализу заготовки с целью выбора подходящей.

Научная новизна:

- В разработке и применении алгоритма анализа заготовок с помощью камеры и библиотеки OpenCV;
- В разработке алгоритма поиска заготовки в кадре с помощью алгоритмов машинного обучения.

Достоверность и обоснованность полученных результатов магистерской диссертации достигнута в ходе тестовых испытаний на оборудовании предприятия ООО «Амурсталь».

Практическая значимость. Практическая значимость и ценность полученных результатов исследования заключается:

- в уменьшении затрат на более дорогостоящее и менее универсальное оборудование для анализа заготовок;
- уменьшении вероятности заполнения документов неверными данными о заготовке за счет автоматизация заполнения характеристик заготовки в документации (например, для сталелитейного производства - паспорте плавки).

Личный вклад автора заключается в анализе применения технологий компьютерного зрения, в применении математической модели приложения на основе изученных технологий разработки и алгоритмов анализа и отслеживания объектов, в разработке программного обеспечения на основе машинного обучения и компьютерного зрения, а также тестировании разработанного приложения в ООО «Амурсталь».

Основные положения, выносимые на защиту:

- Алгоритм определения заготовки в кадре;
- Алгоритм получения размеров найденной заготовки в кадре;
- Разработанный программно-аппаратный комплекс.

Апробация результатов исследования. Основные результаты работы докладывались и обсуждались на следующих семинарах и научных конференциях:

Карамышев, А. Н. Петрова // Международная научно-практическая конференция молодых учёных «Молодёжь и наука: актуальные проблемы фундаментальных и прикладных исследований», 10-14 апреля 2023 г.: в 4 ч. / редкол. : Э.А. Дмитриев (отв. ред.) [и др.]. – Комсомольск-на-Амуре : ФГБОУ ВО «КНАГУ», 2023. – С. 228-230.

Объём и структура диссертационной работы. Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников и приложения. Объем диссертации составляет 80 страниц. Текст работы содержит 4 таблицы и 37 рисунков. Список литературы включает 28 источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность работы, а также приводится общая характеристика работы, включающая в себя следующие структурные элементы:

- цель исследования;
- задачи исследования;
- объект и предмет исследования;
- методы исследования;
- новизна полученных результатов;
- достоверность и обоснованность результатов исследования;
- практическая значимость и ценность работы;
- личный вклад автора;
- основные положения работы, выносимые на защиту;
- апробация результатов работы;
- публикации;
- структура и объем МД.

В первой главе рассматриваются алгоритмы компьютерного зрения для измерения заготовки и алгоритмы глубокого машинного обучения для поиска, сегментации и отслеживания заготовки. Проведён анализ систем поиска объектов в кадре. Рассмотрены основные технические и программные средства для решения задач компьютерного зрения.

Распознавание образов относится к одной из важнейших задач искусственного интеллекта. Она связана со множеством различных областей исследований. связи с этим оптимизация систем по обнаружению объектов на изображениях и поиск новых решений по классификации исследуемых объектов является актуальной задачей в настоящее время. Для построения собственной системы компьютерного зрения необходимо применить основы обработки изображений, выбрать методику детектирования и распознавания объектов, а также программно реализовать получение данных с помощью камер глубинного зрения. Глава состоит из 7 разделов.

Анализ камер глубины. Для реализации компьютерного зрения была установлена необходимость применения специализированного оборудования – камеры глубинного зрения.

Способы получения 3D-изображений в двумерном формате на основе нескольких базовых технологий и применении камер машинного зрения известны уже достаточно давно и традиционно использовались в научной и военной сферах. Но современное стремительное развитие и рост аппаратных возможностей в области цифровой обработки значительно ускорили внедрение таких систем в сферу автоматизации процессов контроля и систем безопасности, в том числе и производственной.

Для разработки приложения по измерению заготовок необходима камера с высоким разрешением и большим углом обзора, потому что длина измеряемых заготовок может достигать 11 метров. Исходя из анализа наиболее подходящей камерой является Intel RealSense D415.

Анализ библиотек компьютерного зрения. В настоящее время количество новых решений, актуальных приложений и библиотек для компьютерного зрения

стремительно возрастает и, хотя некоторые из них все еще находятся на стадии фундаментальных исследований, уже достаточное их количество применяется в коммерческих продуктах, где они органично встраиваются в более крупную систему, которая способна решать сложные задачи, например, в области медицины или в процессе измерения и контроля качества на производствах.

Среди различного числа библиотек наиболее предпочтительной библиотекой для решения задачи по анализу заготовок является библиотека OpenCV так как она имеет наиболее широкий инструментарий по преобразованию изображений.

OpenCV располагает набором алгоритмов для компьютерного видения, обработки изображения, а также имеет встроенные многочисленные алгоритмы общего назначения с открытым кодом, позволяющие реализовать значительную часть математических операций над изображениями. Библиотека может быть использована на наиболее популярных на сегодняшний день платформах - Windows, Linux, Mac, Android, IOS. OpenCV имеет модульную структуру. Это означает, что пакет включает отдельные динамические и статические библиотеки.

Подготовка набора данных для обучения. Подготовка набора данных является одним из самых важных этапов при подготовке к обучению, плохо размеченный и подготовленный дата сет может стать причиной низких результатов обучения. Поэтому специалистами машинного обучения, на этапе подготовки были выявлены наиболее оптимальные методы при обработке данных, а специально разработанные инструменты позволяют упростить и сократить работу по разметке.

Задача сегментации. Ключевой задачей при разработке системы оценки заготовок является обнаружение движущейся заготовки в кадре, а также её дальнейшая сегментация с целью её измерения. Сегментация изображения, это процесс разделения изображения на несколько сегментов. В этом процессе каждый пиксель изображения связан с типом объекта. Существует два основных типа сегментации изображений – семантическая сегментация и сегментация экземпляров.

Готового решения, которое позволит удовлетворить всем требованиям, предъявляемым нашим предприятием, нет. Все предлагаемые решения требуют серьезной доработки и адаптации для работы в условиях нашего предприятия.

Для достижения поставленной цели необходимо решить следующие задачи:

1) исследовать принципы работы библиотеки Pyrealsense и на её основе разработать алгоритм получения трехмерных координат контуров, найденных с помощью машинного обучения, объектов;

2) изучить архитектуру сегментации изображений U-net и на её основе обучить нейронную сеть по выделению заготовки в кадре.

3) Разработать приложение, использующее библиотеку Pyrealsense и нейронную сеть, основанную на архитектуре U-net для измерения заготовки в кадре.

Во второй главе рассматривается применение алгоритмов компьютерного зрения и машинного обучения для решения поставленной задачи.

Применение камеры глубины с помощью библиотеки pyrealsense. При длительной работе камеры RealSense могут возникать условия, которые со временем приводят к деградации производительности, например, воздействие экстремальных температурных циклов или чрезмерных ударов и вибрации. Intel предоставляет набор инструментов для повторной калибровки камер до исходного заводского состояния и сохранения наилучшей производительности в течение всего срока службы. Эти инструменты включают «Self-Calibration» (само калибровку), «Dynamic Calibration» (Динамическую калибровку) и «OEM calibration» (Калибровку OEM).

В рамках разработки системы измерения заготовок использовалась само калибровка. Само калибровка состоит из двух частей, а именно «On-Chip Calibration» и «Tare Calibration». «On-Chip Calibration» предназначена для минимизации шума. «Tare Calibration» предназначена для повышения точности определения глубины.

После проведения калибровки камеры и убедившись в корректности работы камеры её можно использовать для работы в системе измерения заготовок. Алгоритм измерения объектов с помощью камеры состоит из следующих шагов:

1 На вход получаем контур необходимого нам объекта, определенный с помощью обученной нейронной сети;

2 С помощью библиотеки OpenCV определяем координаты точек на концах найденного контура для вычисления длины объекта;

3 С помощью библиотеки pyrealsense и камеры глубины находим получаем глубину кадра для точек на концах контура;

4 Зная координаты этих точек в трехмерном пространстве посчитать расстояние между ними, которое и будет являться длиной объекта.

Для корректного получения координат необходимо использовать внутренние параметры камеры точность, которых была улучшена в результате калибровки. Расчеты горизонтальной и вертикальной координат происходят по формулам:

$$x = (n - prx) * \frac{z}{fx},$$

$$y = (m - pry) * \frac{z}{fy},$$

где n – горизонтальная координата точки полученная из RGB-модуля,

m – вертикальная координата точки полученная из RGB-модуля,

prx – фокусное расстояние плоскости изображения, кратное ширине пикселя,

pry – фокусное расстояние плоскости изображения, кратное высоте пикселя,

fx – горизонтальная координата главной точки изображения в виде смещения пикселя от левого края,

fy – вертикальная координата главной точки изображения в виде смещения пикселя от верхнего края,

z – глубина кадра в точке с горизонтальными координатами (n, m).

Расчет глубины кадра происходит с помощью стандартных инструментов библиотеки pyrealsense.

Семантическая сегментация. Сегментация – это процесс разделения изображения на несколько областей. Пиксели этих областей обычно должны обладать определенными характеристиками. Чаще всего сегментация изображения используется для захвата более четких границ объекта.

Для решения задачи по выделению заготовки в кадре необходимо воспользоваться семантической сегментацией. Семантическая сегментация изображений

– это задача разделения частей изображения на подгруппы пикселей, принадлежащих к соответствующим объектам, с его классификацией.

Одной из наиболее популярных архитектур для построения нейронных сетей является архитектура U-net рисунок 1.

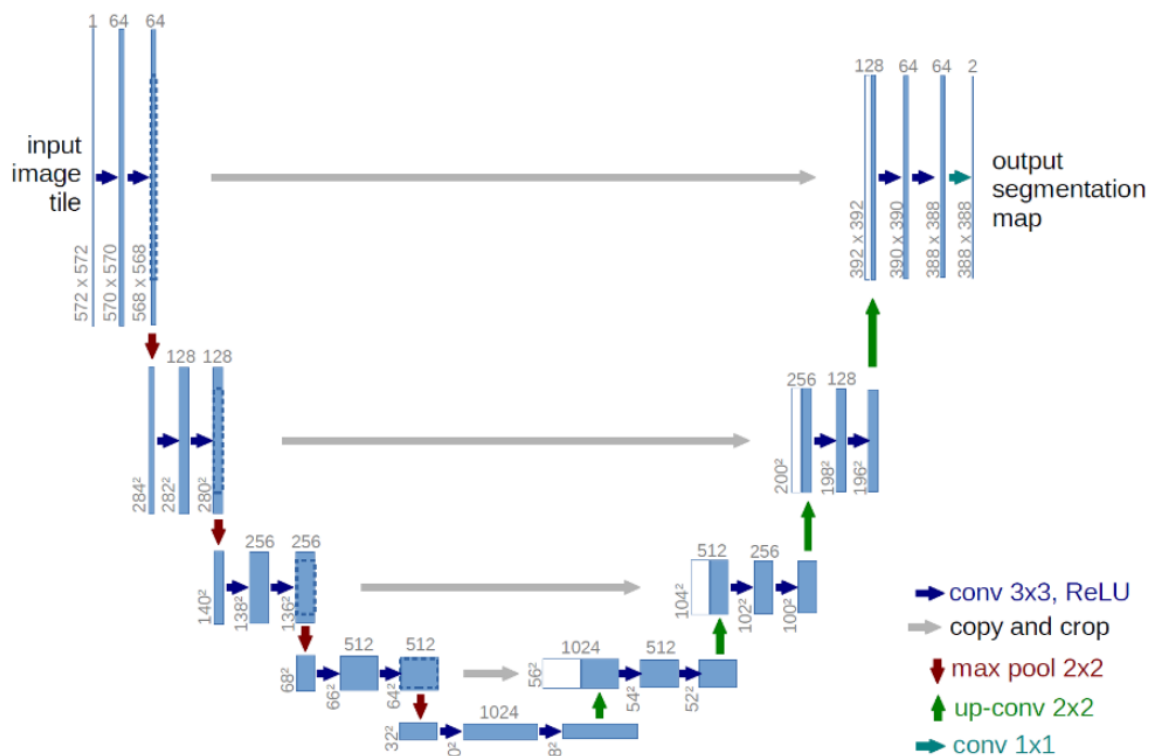


Рисунок 1 – Архитектура U-net

Сеть обучается методом стохастического градиентного спуска на основе входных изображений и соответствующих им карт сегментации. Из-за свертки выходное изображение меньше входного сигнала на постоянную ширину границы. Применяемая попиксельно, функция softmax вычисляет энергию по окончательной карте свойств вместе с функцией кросс-энтропии.

Ключевой особенностью U-net является применение механизма skip-connections. Как следует из названия, skip-connections в глубокой архитектуре обходят некоторые слои нейронной сети и передают выходные данные одного слоя в качестве входных данных для следующих уровней. Это стандартный модуль, представляющий альтернативный путь для градиента с обратным распространением.

Реализация нейронной сети на основе архитектуры U-NET. Алгоритм, реализующий работу нейронной сети, основанной на архитектуре U-net представлен на рисунке 2.

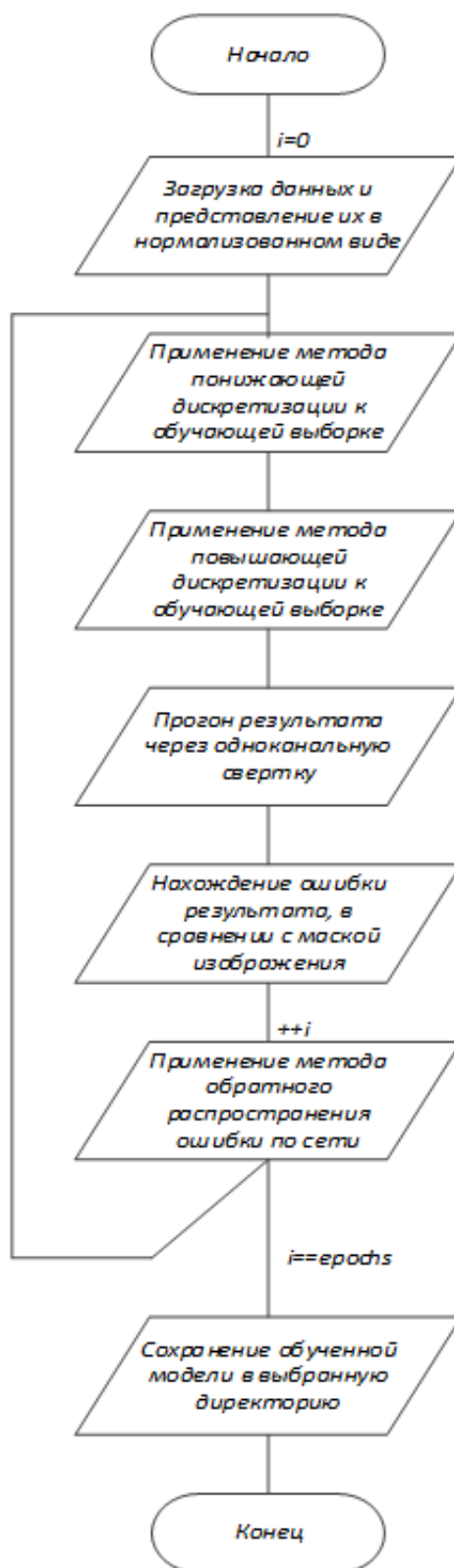


Рисунок 2 – Алгоритм работы нейронной сети

Для решения данной задачи сегментации были выбраны следующие параметры модели нейронной сети, построенной на основе архитектуры нейронной сети U-net:

- Размер обучающей выборки – 1000 изображений;
- Функция ошибки бинарная кросс-энтропия;
- Оптимизатор Адам;
- Сжатие изображений на входе до 512x512x3, где 512x512 – ширина и высота изображения, а 3 – число каналов в изображении;
- Число каналов на каждом этапе 3-64-128-256-512-512-256-128-64-3-1.

Таким образом, в результате обучения была получена нейронная сеть для выделения объекта в кадре. Благодаря правильно подобранным параметрам обученная нейронная сеть получилась достаточно точной для её внедрения в систему измерения заготовок.

В роли функции ошибки была выбрана бинарная кросс-энтропия. Данная функция ошибки выбрана в связи с тем, что изображения, образующееся на выходе нейронной сети и маска, соответствующего на выходе изображения, представляют собой бинарные (так как в данной модели реализована нормализация обучающей выборки) одноканальные изображения. Исходя из этого, бинарная кросс-энтропия, как функция ошибки, подходит для решения данной задачи лучше всего.

В третьей главе магистерской работы описывается практическая реализация программного обеспечения. Данный раздел содержит подробную инструкцию по работе с разработанным приложением.

Основным модулем системы является модуль анализа заготовки. Интерфейс данного окна с включенной модулем анализа заготовки представлен на рисунке 4. В случае, если камера не будет найдена пользователь увидит сообщение об ошибке. Если камера подключена пользователь увидит изображение с камеры. Обнаруженные заготовки и их параметры записываются в базу данных

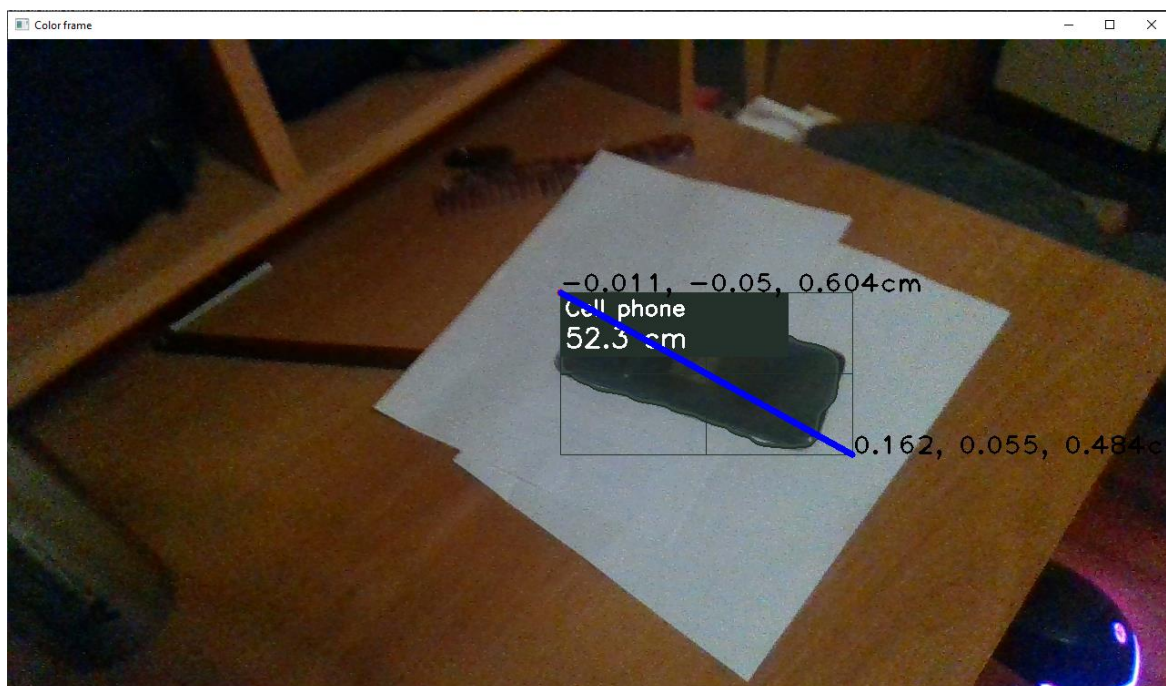


Рисунок 3 –Окно «Анализ заготовки» с работающей камерой

Камера RealSense неспособна обеспечивать качественное измерение заготовок в промышленных условиях. Для повышения точности измерения и отказоустойчивости системы необходимо использовать промышленную камеру. На данный момент рынок камер глубины активно развивается и расширяется, из-за чего возникает необходимость в возможности быстрого внедрения камер глубины различных моделей и от разных производителей, чтобы иметь возможность в кратчайшие сроки сменить камеру на более быструю и точную. Поэтому одной из ключевых задач является внедрение возможности динамического подключения между различными камерами глубины, а также внедрение возможности динамической настройки этих камер.

Другой задачей является получение других параметров, заготовки, а именно:

- 1 измерение скорости движения заготовки;
- 2 анализ заготовки на наличие крупных дефектов;
- 3 анализ корректности движения заготовок по конвейеру.

В заключении перечислены основные результаты работы.

В приложении приведен исходный код программы.

Список использованных источников

- 1 Ермилов, Д. Ю. Обнаружение объектов на основе алгоритмов глубокого обучения / Д. Ю. Ермилов // Современное состояние, проблемы и перспективы развития отраслевой науки: Материалы Всероссийской конференции с международным участием, Москва, 01–05 февраля 2019 года. – Москва: Издательство «Перо», 2019. – С. 402-405. – EDN TQUBCF;
- 2 Венцов, Н. Н. Гибридный подход к обнаружению объектов интереса в видеопотоке / Н. Н. Венцов, Л. А. Подколзина // Фундаментальные основы физики, химии и механики наукоёмких технологических систем формообразования и сборки изделий: Сборник трудов научного симпозиума технологов-машиностроителей, Ростов-на-Дону, 22–26 сентября 2020 года. – Ростов-на-Дону: Донской государственный технический университет, 2020. – С. 289-296. – EDN MYSYFZ;
- 3 Сергиенко С. С. Метод решения задачи компьютерного зрения по поиску простых форм на изображении с использованием алгоритма Хафа / С. С. Сергиенко, А. В. Федорова // Актуальные проблемы инфотелекоммуникаций в науке и образовании: сборник научных статей: в 4х томах, Санкт-Петербург, 24–25 февраля 2021 года. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2021. – С. 404-408. – EDN MGBMNC;
- 4 Ледяев В. П. Обнаружение движущихся объектов с помощью компьютерного зрения и библиотеки opencv / В. П. Ледяев // Экономика и социум. – 2017. – № 12(43). – С. 1943-1947. – EDN YRFLMV.
- 5 Дэвид Форсайт, Жан Понс. Компьютерное зрение. Современный подход Computer Vision: A Modern Approach. – М.: «Вильямс», 2004. «Компьютерное зрение: технологии, рынок, перспективы», отчет TAdviser, 2018.43.
- 6 Goodfellow, Y. Bengio, A. Courville. «Deep Learning Book», MIT Press, 2016 <http://deeplearningbook.org>.
- 7 M. Nielsen, «Neural Networks and Deep Learning», Determintation Press, 2015 <http://neuralnetworksanddeeplearning.com>.

8 W. Dai, C. Dai, S. Qu, J. Li, S. Das «Very Deep Convolutional Neural Networks for Raw Waveforms», ICASSP, 2017 <https://arxiv.org/abs/1610.00087>.

9 J. Yosinski «Understanding Neural Networks Through Deep Visualization», ICML DL Workshop, 2015 <http://yosinski.com/deepvis>.

10 Л. Шапиро, Дж. Стокман. Компьютерное зрение Computer Vision. – М.: Бином. Лаборатория знаний, 2006.

11 Свичкарев, Д. С. Архитектура систем распознавания объектов на изображениях / Д. С. Свичкарев // Вестник современных исследований. – 2018. – № 12.10(27). – С. 345-349. – EDN YVNCLJ

12 Максимов, А. Е. Инновационные образовательные технологии: технология Intel Realsense / А. Е. Максимов, А. А. Акифьев, А. В. Каменский // Современное образование: качество образования и актуальные проблемы современной высшей школы: материалы международной научно-методической конференции, Томск, 31 января – 01 2019 года. – Томск: Томский государственный университет систем управления и радиоэлектроники, 2019. – С. 68-69. – EDN RWLEQE.

13 Лихошерстный, А. Ю. Нейросетевое прецедентное распознавание объектов на аэрокосмических изображениях с использованием частотных признаков / А. Ю. Лихошерстный // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. – 2012. – № 19(138). – С. 136-140. – EDN RMUGKJ.