

## Применение сверточных нейронных сетей для контроля показателей технологического потока производства гранулированных киселей



А. В. Шафрай<sup>1,\*</sup>, А. М. Попов<sup>1</sup>, В. С. Косинов<sup>1</sup>,  
И. Ю. Резниченко<sup>2</sup>, О. Н. Бондарчук<sup>2</sup>

<sup>1</sup> Кемеровский государственный университет , Кемерово, Россия

<sup>2</sup> Кузбасский государственный аграрный университет имени В. Н. Полещукова, Кемерово, Россия

Поступила в редакцию: 20.08.2025

Принята после рецензирования: 02.10.2025

Принята к публикации: 04.11.2025

\*e-mail: shafraia@mail.ru

© А. В. Шафрай, А. М. Попов, В. С. Косинов,  
И. Ю. Резниченко, О. Н. Бондарчук, 2025



### Аннотация.

Применение искусственного интеллекта в пищевой промышленности становится актуальным. Существуют разные способы цифровизации контроля показателей технологического потока производства продукции. Например, использование нейронных сетей для определения размеров гранул сухого киселя. Цель исследования – применить сверточные нейронные сети для контроля показателей технологического потока производства гранулированных киселей посредством локализации гранул киселя на изображении.

Одним из важнейших параметров готовых гранул сухого киселя является их размер, который должен находиться в пределах от 2 до 5 мм. Для разработки модели, определяющей крупные гранулы (более 5 мм), требуется собрать набор данных из изображений гранул разного размера. Выбраны модели локализации объектов из фреймворка Detectron2, которые оптимально справляются с задачей. Описаны принципы работы моделей и качественные показатели для оценки результатов обучения моделей.

Для выявления крупных гранул собран и размечен набор данных из изображений гранул разного размера. Наилучших показателей достигала модель R50-FPN. Наибольшее значение принимала метрика AP50, затем шли AP75 и AP. Модели отлично обучились находить объекты на изображении, достаточно хорошо определяли координаты ограничивающей рамки. В собранном наборе данных отсутствовали размеченные объекты для малых (APs) и средних (APm) размеров, т. к. в исследовании сделан акцент на локализации именно больших гранул. Значения метрики API для всех моделей находились на высоком уровне. Таким образом, выбранный подход к обучению и архитектуре нейронной сети оказался оптимальным для данной задачи. На основе обученной модели разработана программа ЭВМ, которая использует сверточные нейронные сети для детекции больших гранул на снимке с продукцией. В дальнейшем она может быть использована на непрерывных производствах для контроля размера готового продукта и его соответствия технологическим параметрам.

**Ключевые слова.** Искусственный интеллект, нейронные сети, локализация, гранулированные продукты, гранулированные кисели

**Финансирование.** Исследование выполнено за счет гранта Российского научного фонда № 25-26-00136, <https://rscf.ru/project/25-26-00136/>

**Для цитирования:** Шафрай А. В., Попов А. М., Косинов В. С., Резниченко И. Ю., Бондарчук О. Н. Применение сверточных нейронных сетей для контроля показателей технологического потока производства гранулированных киселей. Техника и технология пищевых производств. 2025. Т. 55. № 4. С. 845–855. <https://doi.org/10.21603/2074-9414-2025-4-2613>

## Convolutional Neural Networks in Granulated Kissel Production



Anton V. Shafrai<sup>1,\*</sup>, Anatoliy M. Popov<sup>1</sup>, Vitaly S. Kosinov<sup>1</sup>,  
Irina Yu. Reznichenko<sup>2</sup>, Olga N. Bondarchuk<sup>2</sup>

<sup>1</sup> Kemerovo State University Kemerovo, Russia

<sup>2</sup> Kuzbass State Agricultural University, Kemerovo, Russia

Received: 20.08.2025

Revised: 02.10.2025

Accepted: 04.11.2025

\*e-mail: kosinov\_vs@mail.ru

© A.V. Shafrai, A.M. Popov, V.S. Kosinov, I.Yu. Reznichenko,  
O.N. Bondarchuk, 2025



### Abstract.

Artificial intelligence can be used to monitor production parameters in the food industry. Kissel is a jelly-like fruit or berry starch drink. Instant kissel usually consists of granules. Neural networks may help to control the size of kissel granules. In this research, convolutional neural networks monitored the production parameters of granulated kissel powder by localizing granules in an image. Size is the most important parameter of kissel granules: it should remain between 2 and 5 mm. To detect larger granules ( $\geq 5$  mm), the network was provided with a visual dataset of granules of varying sizes. The localization models were developed using Detectron2.

The research yielded a set of optimal operating principles and quality metrics. The R50-FPN model achieved the best results. The AP50 metric had the highest value, followed by AP75 and AP. The models performed well in visual detection and successfully determined the coordinates of the bounding rectangle. The resulting dataset did not label objects for small (APs) and medium (APm) sizes because the study focused on localizing large granules. The API metric values for all models were high. The approach to AI training and neural network architecture proved optimal for food production control. The trained model made it possible to develop a computer program based on convolutional neural networks that demonstrated good results in detecting large granules in instant kissel powder. The new program can be used in continuous production to monitor the size of finished products and their compliance with process parameters.

**Keywords.** Artificial intelligence, neural networks, localization, granulated products, granulated kissel

**Funding.** The study was supported by the Russian Science Foundation, grant No. 25-26-00136, <https://rscf.ru/en/project/25-26-00136/>

**For citation:** Shafrai AV, Popov AM, Kosinov VS, Reznichenko IYu, Bondarchuk ON. Convolutional Neural Networks in Granulated Kissel Production. *Food Processing: Techniques and Technology*. 2025;55(4):845–855. (In Russ.) <https://doi.org/10.21603/2074-9414-2025-4-2613>

### Введение

Применение искусственного интеллекта в промышленности стало одним из ключевых трендов цифровой трансформации различных отраслей. В пищевой промышленности внедрение искусственного интеллекта открывает широкие возможности для автоматизации процессов, повышения точности контроля качества, оптимизации производственных мощностей и моделирования параметров как используемых ресурсов, так и производимой продукции [1–3]. Особенно перспективным направлением является использование нейронных сетей и машинного обучения для анализа больших объемов данных, получаемых в режиме реального времени с датчиков оборудования и линий производства [4]. Такие технологии позволяют не только выявлять отклонения от нормы на ранних стадиях,

но и предсказывать возможные сбои, что способствует снижению брака и увеличению общей рентабельности производства.

В частности, производство гранулированных продуктов, таких как кисели, представляет собой сложный технологический процесс, требующий строгого соблюдения параметров сырья, температурного режима, влажности и других факторов. На сегодняшний день в этой области недостаточно внедрены современные методы искусственного интеллекта, несмотря на растущий запрос со стороны предприятий на автоматизированное управление качеством [5]. Внедрение систем на основе нейросетей позволит осуществлять непрерывный контроль за состоянием продукции, распознавать дефекты гранул по визуальным и физико-химическим характеристикам, а также адаптировать параметры производства

под изменяющиеся условия в автоматическом режиме. Это не только повысит стабильность выпускаемой продукции и ее соответствие стандартам, но и снизит затраты на контроль технологического потока, увеличив экономическую эффективность производства.

Цель исследования – применить сверточные нейронные сети для контроля показателей технологического потока производства гранулированных киселей посредством локализации гранул киселя на изображении. Для выполнения цели определен ряд задач:

1. Подготовить набор данных (датасет) гранул различных размеров.
2. Обучить сверточную нейронную сеть на подготовленном наборе данных.
3. Провести оценку качества работы модели.
4. Разработать программное обеспечение на основе обученной сети.

Таким образом, выполнение указанных задач позволит оценить возможности нейронных сетей, способных локализовать гранулы киселя на изображении, что откроет возможности для дальнейшего анализа их характеристик и повышения общего уровня цифровизации контрольных операций на производстве.

### Объекты и методы исследования

Технологическая схема производства быстрорастворимых гранулированных киселей представлена на рисунке 1. Она включает несколько основных стадий, каждая из которых направлена на обеспечение высокого качества конечного продукта. Одним из наиболее определяющих этапов в этой схеме является процесс гранулирования.

Существует несколько методов гранулирования. Один из наиболее популярных в пищевой промышленности – гранулирование методом окатывания. Оно реализуется в специальных аппаратах (грануляторы-окатыватели). Основные виды таких устройств – тарельчатые и барабанные установки. Принцип их работы заключается в подъеме материала на определенную высоту за счет вращательного движения рабочей поверхности с последующим скатыванием вниз под действием силы тяжести частиц, формирующих гранулы. К ключевым факторам, обуславливающим низкое удельное энергопотребление процесса, относится то, что значительная часть энергии расходуется на пре-

одоление потенциальной энергии, то есть на подъем материала. При скатывании происходят многочисленные взаимодействия между гранулами, частицами исходного материала и внутренними поверхностями аппарата, что способствует уплотнению и росту гранул. Подъем обеспечивается за счет комбинации центробежных сил и сил трения между частицами и поверхностью гранулятора.

Тарельчатый гранулятор представляет собой вращающуюся чашу с бортом, установленную под углом 35–55° к горизонту. Внутри нее материал движется по сложной спиральной траектории. На частицу действуют сила тяжести, центробежная сила, силы трения о дно и борт, а также реакция дна. При достижении определенной высоты частица отрывается от борта и скатывается вниз, образуя восходящую и нисходящую ветви движения. В процессе окатывания наблюдается сегрегация по крупности: мелкие фракции остаются у дна, более крупные – выше, а самые крупные гранулы движутся вблизи борта.

Эффективность тарельчатого гранулятора во многом зависит от того, насколько полно используется площадь дна чаши. Исследования показали [6], что максимальная эффективность достигается, когда мелкие частицы отрываются от борта в самой верхней точке. Это условие выполняется при определенном соотношении частоты вращения, диаметра тарели, углов ее наклона и естественного откоса материала. С увеличением частоты вращения для сохранения оптимального отрыва необходимо увеличивать угол наклона, что, в свою очередь, повышает высоту скатывания и силу удара гранул в нижней точке.

Для предотвращения смещения материала к борту при высоких оборотах и обнажения дна применяют меры по увеличению трения. Например, футеруют тарель резиной или поддерживают слой гарнисажа, представляющего собой тонкий слой материала на дне. Гарнисаж не только повышает сцепление, но и защищает поверхность от абразивного износа, а также выполняет демпфирующую функцию при колебаниях влажности, поглощая избыток жидкости и постепенно ее отдавая. Такой режим особенно важен при мокром гранулировании.

В процессе окатывания происходят существенные изменения в структуре дисперсной системы. Исходная

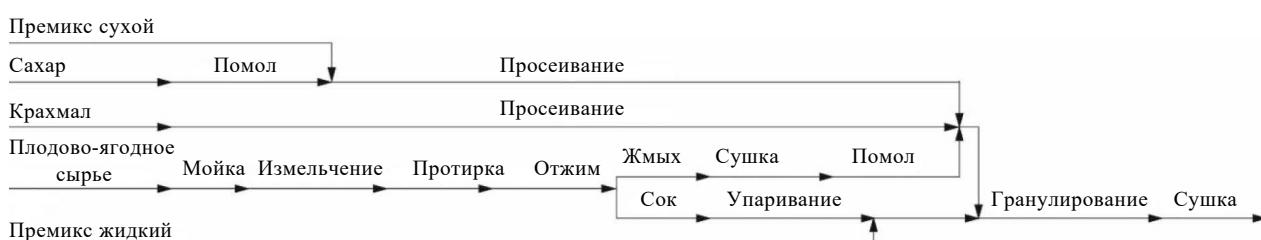


Рисунок 1. Схема технологического потока быстрорасторимых гранулированных киселей

Figure 1. Instant kiszel production

рыхлая масса с трехфазной структурой постепенно переходит в плотную гранулу. При этом жидкую фазу перераспределяется: из пленочного и капиллярно-стыкового состояния она переходит в поровое пространство, формируя капиллярную воду в фуникулярном состоянии. В зонах контакта каркасных зерен концентрируются ультрадисперсные частицы, образуя структурированную суспензию с высокими реологическими свойствами. Эта псевдофаза имеет тиксотропные характеристики и помогает повысить прочность гранул за счет капиллярно-молекулярного механизма связи.

Рост и уплотнение гранул напрямую зависят от скорости капиллярного влагообмена. Поверхность растущей гранулы должна быть слегка переувлажнена, чтобы обеспечить налипание новых частиц. Избыток влаги вытесняется из ядра гранулы как за счет капиллярного потенциала, так и за счет механического выдавливания. На конечной стадии, когда гранула сформирована, происходит релаксация напряжений и выравнивание градиентов плотности и влажности по объему. В это время поровая суспензия переходит в устойчивое коагуляционно-структурное состояние.

Механизм гранулообразования зависит от выбранного способа: сухой, полумокрый или мокрый. Мокрый способ подразумевает, что материал изначально содержит достаточное количество влаги, частицы соединены менисками жидкости и могут перемещаться, образуя более плотную укладку. Зародыши гранул формируются случайным образом из увлажненных комков, которые затем уплотняются. Рост гранул происходит за счет налипания мелких агломератов, разрушающихся в процессе окатывания. Условие присоединения частицы к грануле определяется соотношением ее диаметра, радиуса гранулы и угла трения. При перекатывании гранулы через частицу развивается значительное давление, достигающее сотен атмосфер, что способствует вдавливанию частицы и уплотнению структуры.

При сухом способе влага подается распылением. Капли жидкости вызывают локальное слипание частиц, образуя зародыши. Эти комки оказываются переувлажненными, и при окатывании избыток жидкости вытесняется на поверхность, обеспечивая налипание новых частиц. Однако при наличии растворимых компонентов процесс усложняется. По мере растворения твердой фазы объем жидкой фазы увеличивается, что приводит к размягчению и разрушению гранул. Это свидетельствует о том, что для устойчивого гранулирования важно не только количество влаги, но и ее подвижность в структуре.

Эксперименты показали, что, несмотря на малое содержание, ультрадисперсные фракции играют ключевую роль в формировании прочной структуры. Они концентрируются в точках контакта зерен, образуя малоподвижную суспензию, в то время как влага перемещается в виде раствора. Это означает, что именно динамика жидкой фазы при уплотнении каркаса определяет эффективность процесса.

Таким образом, процесс окатывания представляет собой сложное сочетание механических, капиллярных и реологических явлений, зависящих от свойств материала, режимных и конструкционных параметров оборудования. Гранулометрический состав влажных гранул, получаемых окатыванием, легко регулируется, обеспечивая узкий диапазон размеров при диаметре от 5 до 30 мм за счет технологической классификации на тарели гранулятора. Размер выходящих гранул определяется высотой борта. Указанные размеры гранулята не отвечают концепции нашей работы создания быстрорасторвимых сухих концентратов, исходя из которых размер гранул должен быть 1–3 мм [6].

Технологический поток оценивается по наиболее важным технологическим показателям промежуточных и готового продуктов на разных стадиях производства. Показатели и их допустимые значения определяются экспертами, например, технологами и инженерами производства. В данном исследовании в качестве оцениваемого параметра выбран размер гранул. При производстве киселей количество готовых гранул (больше 5 мм) не должно превышать 10 %. Поэтому далее нейронную сеть обучали для распознавания таких гранул.

Приведем общую последовательность действий для формирования набора данных, выбора метода обучения и построения модели (рис. 2).

1. Создание набора данных, содержащего изображения и их разметку:

- собирается достаточный набор изображений, определяются их размеры и размеры, которые будут использоваться в обучении (т. к. некоторые архитектуры нейронных сетей работают с изображениями определенного размера), изображения форматируются под требуемые размеры;
- полученные изображения размечаются (под разметкой следует понимать присвоение аннотации, метки или тега элементу/ам, находящимся на изображении), присвоенный индикатор и будет являться целевым значением;
- очень часто собрать необходимый объем данных для обучения затруднительно. В таком случае прибегают к аугментации, имеющиеся размеченные изображения проходят набор трансформаций: вращение исходных изображений, инверсия, зеркализование и т. д. Таким образом в разы увеличивается объем датасета, а получаемая модель учится реагировать на различные варианты одних и тех же изображений, тем самым, приобретает лучшие обобщающие качества.

2. Импорт всех необходимых библиотек для работы с датасетом и нейронными сетями.

3. Начало итерационного процесса подбора архитектуры, гиперпараметров и параметров обучения:

- выбирается архитектура нейронной сети, причем выбирается как семейство, так и конкретная модель. Например, у архитектуры ResNet есть несколько модификаций: ResNet18, ResNet32, ResNet50 и т. д. Чаще



Рисунок 2. Принципиальная схема применения нейронных сетей

Figure 2. Neural networks: Application scheme

всего различия между видами одной архитектуры заключаются в количестве скрытых слоев. Поэтому выбирают подходящий вид архитектуры (на основании дальнейшего использования можно сделать вывод: чем больше слоев, тем больше вычислительных мощностей потребуется в конечном использовании), затем из разных семейств берутся схожие по размерам модели для дальнейшего тестирования. Так как в основном используется технология дообучения, модели берутся уже обученные (в данном случае на датасете ImageNet), с готовыми весами. Такие модели уже способны классифицировать изображения на 1000 классов; – последние слои выбранных моделей изменяются под решаемую задачу. Чаще всего готовые модели классифицируют на 1000 заранее определенных классов. Поэтому под конкретную задачу последние слои (они идут полностью связанные) переписываются. Например, если требуется классификация на 2 класса: норма / брак, то необходим один выходной нейрон, который будет предсказывать вероятность брака: 0 – норма, 1 – брак. Далее сеть дообучают на собранном датасете; – подбирают гиперпараметры, в число которых входит функция потерь, оптимизатор, регуляризаторы, приемы обучения, такие как дроп-аут или батч-нормализация; – начинается обучение с подбором параметров (например, размер батчей, количество эпох, размеры тренировочной, валидационной и тестовой выборки и др.);

– подбираются метрики, по которым судят об успешности той или иной комбинации всех переменных и точности обученной модели. Все подпункты пункта 3 обозначают итерационный процесс, который длится до достижения желаемой точности модели, в большинстве случаев представляя из себя полный перебор всех возможных комбинаций.

Для успешного обучения нейронной сети необходимо предварительно собрать и подготовить соответствующий набор данных – датасет. Он представляет собой совокупность изображений, на которых представлены объекты, подлежащие распознаванию моделью, в данном случае – отдельные гранулы киселя.

Снимки гранул сделаны на производственной площадке компании ООО «НПО Здоровое Питание» (Кемерово, Россия), что позволило обеспечить высокую презентабельность данных относительно реальных условий производства. Для сбора изображений использовалась цифровая камера D5300 kit 18-140mm (Nikon, Япония), установленная на этапе сушки гранул после гранулирования. Изображения сделаны при различном освещении и ракурсах, чтобы обучаемая модель могла корректно обрабатывать данные вне зависимости от внешних факторов. На каждом снимке можно наблюдать как отдельные, так и сгруппированные гранулы, различающиеся по размеру, форме и степени однородности поверхности. Это позволяет модели

лучше учитывать вариативность продукта и повышает ее устойчивость к сложным случаям распознавания. Полученные изображения затем объединяли в единый структурированный датасет, который стал основой для дальнейшего этапа – его разметки и последующего обучения нейросетевой модели.

Для получения целевых значений, требуемых для обучения нейронных сетей детектированию крупных гранул, полученные снимки размечали. Для задачи локализации крупных частиц на изображении требуется присвоить метку каждой крупной грануле. Разметка осуществлялась в программе для аннотирования изображений и видео Computer Vision Annotation Tool. Созданная разметка может быть сохранена в различных форматах, используемых для обучения нейронных сетей. В данной работе разметка сохранена в формате COCO 1.0. Целевые значения хранились в файле формата json, содержащего информацию об изображении, включая путь к файлу, размер и список аннотированных объектов.

Нейронные сети представляют собой вычислительные модели, которые состоят из взаимосвязанных узлов, называемых нейронами. Они способны обучаться на основе получаемых данных [7]. Нейронные сети используются при решении различных задач, в число которых входит локализация объектов на изображении. Существует несколько типов нейронных сетей (архитектур нейронных сетей), для работы с изображениями применяются сверточные нейронные сети (СНС). Они способны анализировать не только сами данные, но и их взаимное расположение. Такая особенность необходима при исследовании изображений, потому что они состоят из пикселей. Именно расположение пикселей создает объекты, которые человек воспринимает на изображении.

Кроме того, СНС представляют собой достаточно большие модели по количеству параметров, поэтому их разработка и обучение является очень сложным и затратным процессом, доступным крупным технологическим компаниям и исследовательским организациям. В связи с этим для решения прикладных задач используют уже готовые и обученные модели различных архитектур. Их дообучают под решение конкретных примеров. Такая технология называется *transfer learning*.

Mask R-CNN является одной из самых современных и эффективных моделей СНС, применяемой для задачи локализации (*instance segmentation*) [8]. Mask R-CNN позволяет не только определить класс объекта, но и выделить его на изображении. Модель Mask R-CNN разработана исследовательской группой Facebook как развитие более ранних архитектур, таких как R-CNN, Fast R-CNN и Faster R-CNN [9].

В рамках данного исследования модель Mask R-CNN реализована с помощью фреймворка Detectron2, который базируется на библиотеке PyTorch и предоставляет гибкие инструменты для обучения и тестирования

моделей *instance segmentation*. Detectron2 рассчитан на работу с RGB-изображениями в традиционных форматах, таких как png или jpg.

Работа модели Mask R-CNN строится на двух основных этапах. На первом этапе модель анализирует изображение и предлагает участки, которые потенциально содержат объекты. Этот процесс осуществляется с помощью сети предложений регионов (RPN) [10]. На втором этапе каждый из участков детально обрабатывается: модель определяет, к какому классу он относится, уточняет границы объекта и строит точную маску его формы. Таким образом, Mask R-CNN объединяет в себе сразу три задачи: обнаружение, классификацию и сегментацию.

Основной частью модели, получающей изображение и извлекающей из него признаки, является *backbone-сеть*. В Detectron2 представлено большое количество *backbone-сетей*. В текущей работе протестированы несколько модификаций *backbone-сети* ResNet50: R50-C4, R50-DC5 и R50-FPN. Данные модификации *backbone-сети* имеют наилучшие показатели среди всех моделей, представленных в Detectron2.

Модель R50-C4 использует выход только четвертого уровня ResNet50 (C4) в качестве финальной карты признаков (рис. 3). Такой подход обеспечивает высокую скорость работы и относительно низкие требования к вычислительным ресурсам, что делает ее хорошим выбором для задач, где важна производительность. Модель R50-DC5 является модификацией R50-C4, в которой последний уровень сети (C5) адаптирован с помощью дилатационных сверточных ячеек для сохранения более высокого пространственного разрешения (рис. 3). Это позволяет улучшить качество обнаружения мелких объектов.

Наиболее продвинутым решением является модель R50-FPN, в которой используется сеть пирамидальных признаков (FPN) [11]. Данный механизм объединяет информацию со всех уровней ResNet50 (C2–C5). FPN строит пирамиду P2–P5, где каждый уровень отвечает за определенный диапазон масштабов объектов (рис. 3). Это позволяет эффективно работать как с крупными, так и с мелкими объектами, обеспечивая наивысшую точность среди перечисленных моделей.

После того как *backbone* извлекает признаки изображения, данные передаются в RPN, которая генерирует *anchor box*. Это набор прямоугольников, внутри которых потенциально могут находиться целевые объекты. Создаются *anchor box* разных размеров и пропорций, чтобы охватить как можно больше возможных вариантов расположения объектов на изображении. Для каждого такого региона модель рассчитывает вероятность наличия в нем объекта и корректирует координаты, чтобы точнее его локализовать.

Полученные таким образом области интереса (ROI) далее подвергаются процессу выравнивания (ROI Align). Этот шаг необходим для того, чтобы привести все найденные области к единому размеру, сохранив

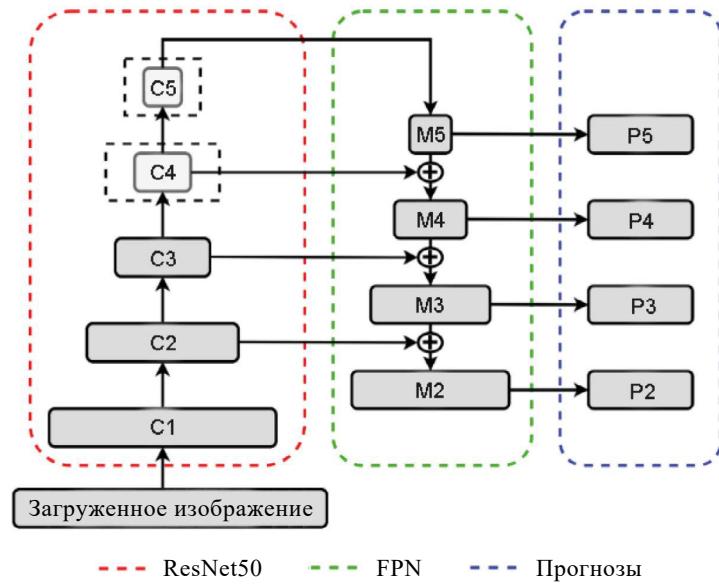


Рисунок 3. Вид backbone-сети, показывающий комбинацию архитектуры ResNet50 с сетью пирамидальных признаков (FPN) и прогнозами

Figure 3. ResNet50 architecture with pyramidal feature network (FPN) and forecasts: Backbone network view

при этом точное пространственное положение объекта. В отличие от более раннего метода ROI Pooling, ROI Align использует билинейную интерполяцию, которая не теряет информацию о расположении пикселей, что особенно важно при построении точных масок сегментации.

После выравнивания каждую ROI приводят к стандартному размеру с помощью операции ROI Align, затем направляют сразу в три параллельных выхода: для определения класса объекта, уточнения координат ограничивающих рамок объекта и построения маски. Все эти шаги происходят одновременно, что делает модель достаточно эффективной даже на сложных изображениях. Таким образом, Mask R-CNN объединяет в одном алгоритме сразу несколько задач: обнаружение, классификацию, а также сегментацию. Описанная организация позволяет модели не только находить объекты, но и точно выделять их форму, даже в том случае, если они частично перекрываются другими объектами.

В процессе обучения модель оптимизирует несколько целевых функций, отвечающих за разные аспекты задачи сегментации. Общая функция потерь рассчитывается по формуле (1) [12]:

$$L = L_{\text{mask}} + L_{\text{cls}} + L_{\text{bbox}} \quad (1)$$

где  $L_{\text{mask}}$  – функция потерь маски;  $L_{\text{cls}}$  – функция потерь классификации;  $L_{\text{bbox}}$  – функция потерь регрессии координат ограничивающей рамки.

Функции потерь маски и классификации рассчитываются по формуле (2):

$$L_{\text{mask}} \text{ и } L_{\text{cls}} = \sum_{i=1}^N y_i \times \log(p(y_i)) + (1 - y_i) \times \log(1 - p(y_i)) \quad (2)$$

Функция потерь регрессии координат ограничивающей рамки рассчитывается по формуле (3):

$$L_{\text{bbox}} = \sum_{i=1}^N |y_{\text{true}} - y_{\text{pred}}| \quad (3)$$

Модель Mask R-CNN обучается таким образом, чтобы минимизировать ошибки сразу по всем компонентам, это позволяет ей одновременно улучшать качество как обнаружения объектов, так и их детального описания.

Еще одним важным этапом в процессе обучения нейронной сети является выбор подходящих гиперпараметров. От их настройки зависит не только скорость сходимости, но и конечное качество модели. В качестве оптимизатора выбрали стандартную настройку Adam с начальной скоростью обучения, равной 0,001. Эта скорость уменьшалась в десять раз после 1000 итераций, чтобы помочь модели лучше сойтись. Также использовался параметр momentum, равный 0,9, который помогал ускорить обучение в нужном направлении. Для каждого изображения обрабатывалось по 128 ROI, чтобы натренировать модель достаточно точно оценивать объекты разного размера. Обучение проводилось в течение 5000 эпох, при этом на каждом шаге отслеживалось значение функции потерь на валидационной выборке для своевременной остановки обучения, как только модель переставала улучшаться. Для более стабильной начальной фазы обучения приме-

нялось 1000 итераций warm-up – на этом этапе скорость обучения постепенно увеличивалась от очень малого значения (0,001) до установленного уровня. Такой подход помог моделям легче адаптироваться к данным и избежать резких скачков весов на ранних этапах.

Расширение датасета за счет аугментации – важный этап при обучении модели нейронной сети. Она представляет собой применение различных преобразований к изображению и его аннотации. Данный подход важен при работе с малыми наборами данных. К изображениям в собранном датасете применяли следующие преобразования:

- RandomFlip – случайное отражение изображения по горизонтали или вертикали;
- RandomBrightness – случайное изменение яркости;
- RandomContrast – случайное изменение контраста;
- RandomSaturation – случайное изменение насыщенности цветов.

Эти преобразования позволяли модели лучше обобщать данные и быть устойчивой к различиям в освещенности и размерах объектов, которые могут встречаться в реальных условиях.

Как упоминалось ранее, для работы использовали набор данных, представленный в формате COCO 1.0 – стандартизированном и широко применяемом формате для задач компьютерного зрения. Одно из ключевых преимуществ этого формата – наличие единой системы оценки качества моделей, что обуславливает корректность сравнения различных архитектурных и подходов.

Для оценки эффективности СНС в рамках COCO используется несколько стандартных метрик, основной из которых является средняя точность (AP) [13]. В отличие от традиционных методов, где используется фиксированный порог перекрытия предсказанной области с истинной (IoU) [14], в COCO вычисляется усредненное значение AP по нескольким порогам IoU (от 0,50 до 0,95 с шагом 0,05). Данное обстоятельство делает оценку более строгой и объективной, т. к. модель должна показывать стабильно высокие результаты на разных уровнях точности локализации.

Кроме общей метрики AP, также рассматриваются следующие показатели:

- AP50 – значение средней точности при IoU = 0,50;
- AP75 – значение средней точности при IoU = 0,75.

На основе приведенных метрик можно оценить работу модели при разной степени строгости к положению и размеру предсказанных ограничивающих рамок объекта.

Дополнительно проводили анализ качества модели в зависимости от размера распознаваемых объектов:

- APs – средняя точность для маленьких объектов (менее 32<sup>2</sup> пикселей);
- APm – средняя точность для объектов среднего размера (от 32<sup>2</sup> до 96<sup>2</sup> пикселей);
- API – средняя точность для больших объектов (более 96<sup>2</sup> пикселей).

Такая детализация давала понимание об эффективности модели в распознавании объектов различного размера, что особенно важно в задачах контроля качества, где необходима точность определения как мелких, так и крупных частиц.

Для обучения модели определили комплекс оборудования и программного обеспечения:

- операционная система: Windows 10 (версия 10.0.19041 SP0) – из-за стабильности и удобства работы;
- интерпретатор: Python 3.8.12, ввиду широкой поддержки библиотек машинного и глубокого обучения и устойчивости версии;
- среда разработки: облачная платформа Google Colab, обеспечивающая доступ к GPU, облачному хранению данных и удобной работе с Jupyter-блокнотами.

Выбранный комплекс позволил эффективно реализовать процесс обучения нейросетевой модели.

## Результаты и их обсуждение

Сделано 30 снимков, содержащих как крупные гранулы (более 5 мм), так и мелкие (менее 5 мм). Размер полученных изображений составил 4320×3240 пикселей. Данные снимки разбивали на 16 частей равного размера. Таким образом, объем изображений в наборе обучающих данных увеличился до 480. Размер изображений в данном случае стал равен 810×810 пикселей. Позже изображения сжимали до размера 512×512 пикселей. Далее размещали 480 изображений, содержащих гранулы большого и малого размеров (рис. 4). Большим гранулам присваивали тег bigGranule.

После завершения процесса обучения сверточной нейронной сети (СНС) получены и проанализированы количественные показатели ее точности. Основными метриками оценки качества модели стали различные значения средней точности (AP), которые дают комплексное представление об эффективности обнаружения гранул на изображении. Результаты, отражающие значение метрик при различных порогах перекрытия



Рисунок 4. Пример разметки больших гранул

Figure 4. Marking larger granules

предсказанной области с истинной (IoU) и размерах объектов, представлены в таблице.

Наилучших показателей достигала модель R50-FPN. Для каждой модели наибольшее значение принимала метрика AP50, затем шли AP75 и AP. Это говорит о том, что модели отлично обучились находить объекты на изображении и достаточно хорошо определяют координаты ограничивающей рамки. Исходя из того, что в исследовании был сделан акцент на локализации гранул именно большого размера, было ожидаемо, что в результате обучения будут отсутствовать значения показателей средней точности для малых (APs) и средних (APm) объектов. Как следствие, модель не обучалась локализации объектов такого размера. Значения метрики API для всех моделей находились на высоком уровне. Это указывает на то, что выбранный подход к обучению и архитектуре нейронной сети оказался оптимальным для данной задачи.

Обученная модель сохранена, для нее разработан интерфейс, позволяющий пользователю загружать и исследовать снимки готовых гранул. Программный код зарегистрировали в качестве программы ЭВМ № 2025680149.

Программа принимает на вход изображение с гранулами. После обработки на выходе формируется

размеченное изображение, на котором каждая обнаруженная гранула выделена рамкой. Цвет рамки соответствует отдельной грануле, рядом выводится текстовая метка и значение вероятности, с которой модель отнесла объект к данному классу. Программа работает стабильно и быстро (до 5 секунд на анализ изображения). Она способна обрабатывать изображения в различных форматах. Предусмотрена возможность работы как с отдельными снимками, так и с несколькими изображениями. Это делает ее удобной для использования в промышленных условиях, где требуется анализ большого количества данных. Обработка выполняется последовательно: сначала изображение загружается; проходит этап предварительной обработки (нормализация, изменение размера, устранение шумов); запускается модель с загруженными весами, которая выполняет детекцию и классификацию объектов. Результаты визуализируются и сохраняются в указанную директорию.

Для наглядности представлены блок-схема алгоритма и интерфейс программы. Блок-схема отражает последовательность всех этапов обработки: от загрузки изображения до вывода размеченного результата. Каждый блок соответствует определенному этапу работы (рис. 5).

Таблица. Значения метрик после процедуры обучения

Table. Metric values after training

	AP	AP50	AP75	APs	APm	API
R50-C4	80,877	83,236	82,111	–	–	80,886
R50-DC5	81,541	84,112	82,556	–	–	81,590
R50-FPN	81,908	84,328	83,165	–	–	81,237

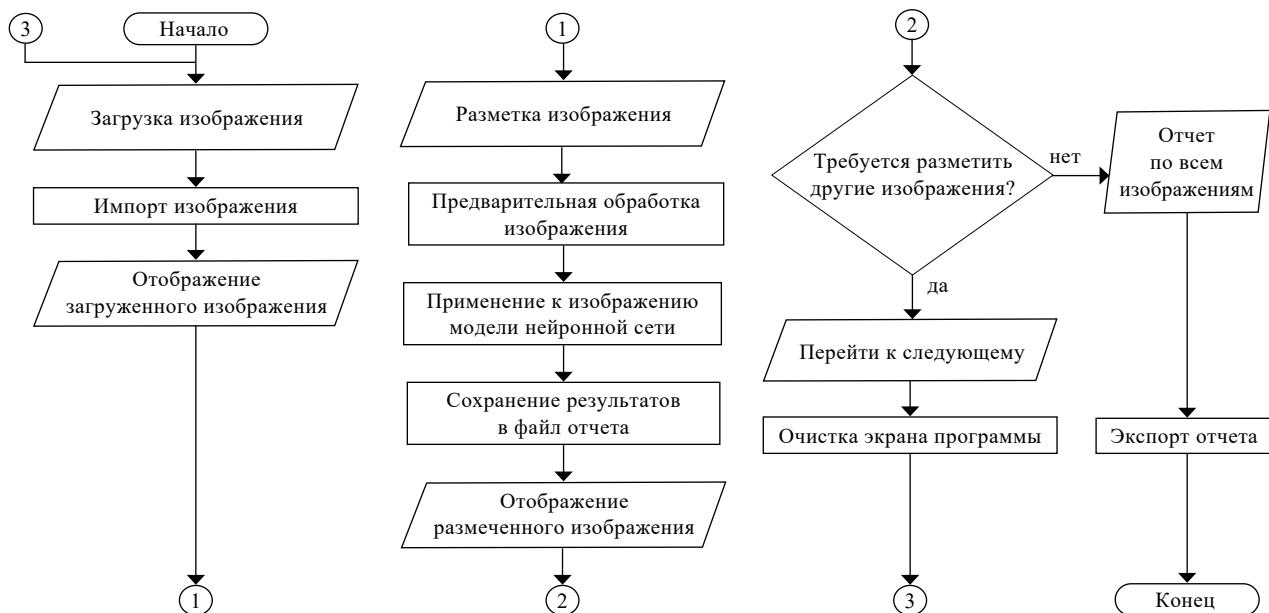


Рисунок 5. Алгоритм работы программы

Figure 5. Operation algorithm: Block diagram

Интерфейс программы, представленный на рисунке 6, выполнен в виде окна с минимальным набором элементов.

Пользователь может выбрать файл, нажав на кнопку «загрузить изображение». После чего выбранное изображение отобразится слева. Далее можно запустить анализ изображения, нажав «разметить изображение». Обработанное изображение отобразится справа. Результаты обработки сохраняются в отдельный файл с отчетом. Если пользователь хочет разметить другие изображения, то для этого требуется выбрать «перейти к следующему». В результате чего поля с изображениями очищаются, и пользователь сможет повторить процедуру для другого изображения. Генерация отчета для любого количества рассмотренных пользователем изображений происходит после нажатия кнопки «отчет по всем изображениям». Интерфейс довольно простой, интуитивно понятный, не требует специальных знаний для эксплуатации.

## Выводы

В ходе проделанной работы собран и подготовлен обучающий набор данных, состоящий из изображений различных типов гранул. Все изображения размечены вручную – на каждом изображении определены гранулы, превышающие допустимый размер, что позволило обучить модель распознавать их самостоятельно.

Для решения задачи разработки модели выбрали одну из современных моделей сверточных нейронных сетей (СНС), подходящую для задачи локализации объектов на изображении. Рассмотрены три моди-

фикации данной модели, которые прошли несколько этапов обучения, включая подбор оптимальных гиперпараметров, аугментацию данных для улучшения обобщающей способности и финальное тестирование на отдельной выборке.

Результаты показали высокую точность модели: она успешно распознавала гранулы большого размера, наличие которых может указывать на нарушения в процессе производства гранулированного киселя. Таким образом, СНС могут использоваться в процессе производства гранулированного киселя для контроля технологического потока.

На основе обученной модели разработана программа ЭВМ, которая использует СНС для детекции больших гранул на снимке с продукцией. В дальнейшем она может оказать пользу на непрерывных производствах для контроля размера готового продукта и его соответствия технологическим параметрам. В таком случае на входе программе сразу будет передаваться видеопоток, и потребуется некоторая перестройка архитектуры. Данный аспект является задачей автоматизации производства, поэтому не рассматривался в данном исследовании.

## Критерии авторства

Авторы в равной степени участвовали в написании рукописи.

## Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

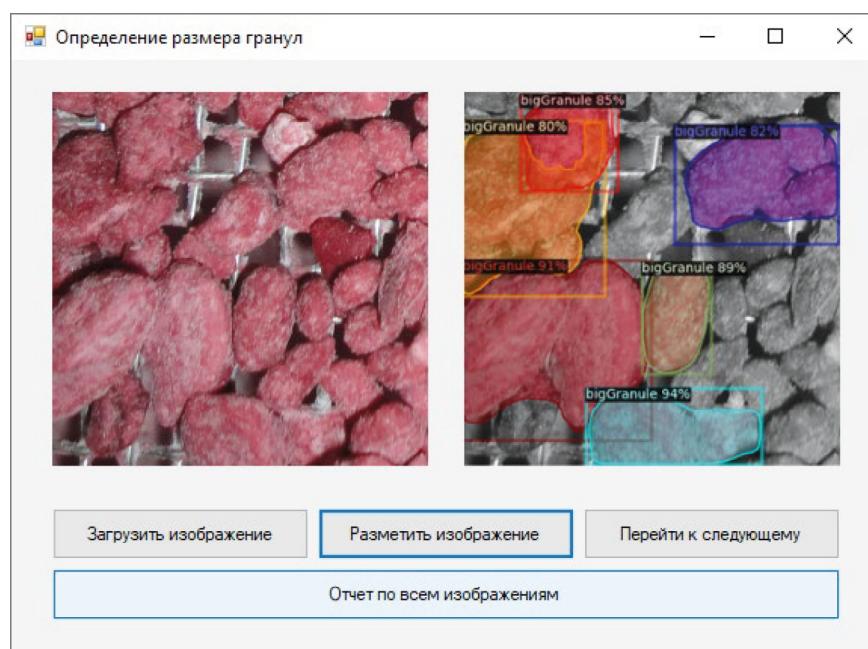


Рисунок 6. Интерфейс разработанной программы

Figure 6. Program interface

### Contribution

The authors contributed equally to the research and bear equal responsibility for the information published in this article.

### Conflict of interest

The authors declared no potential conflict of interest regarding the research, authorship, and/or publication of this article.

### Список литературы / References

1. Shafrai AV, Permyakova LV, Borodulin DM, Sergeeva IY. Modeling the physiological parameters of brewer's yeast during storage with natural zeolite-containing tuffs using artificial neural networks. *Information*. 2022;13(11):529. <https://doi.org/10.3390/info13110529>
2. Королев И. А. Автоматизированное определение дисперсности воздушной фазы в мороженом с применением методов машинного обучения. *Техника и технология пищевых производств*. 2023. Т. 53. № 3. С. 455–464. [Korolev IA. automated measurement of air bubbles dispersion in ice cream using machine learning methods. *Food Processing: Techniques and Technology*. 2023;53(3):455–464. (In Russ.)] <https://doi.org/10.21603/2074-9414-2023-3-2448>
3. Jo DM, Han S-J, Ko S-C, Kim KW, Yang D, *et al.* Application of artificial intelligence in the advancement of sensory evaluation of food products. *Trends in Food Science & Technology*. 2025;165:105283. <https://doi.org/10.1016/j.tifs.2025.105283>
4. Михеев П. Н. Технологии искусственного интеллекта в пищевой промышленности. *Инновации и инвестиции*. 2023. № 4. С. 536–539. [Mikhеев PN. Artificial intelligence technologies in the food industry. *Innovation & Investment*. 2023;(4):536–539. (In Russ.)] <https://elibrary.ru/DYOHTS>
5. Лындина М. И. Пути совершенствования пищеконцентрантной отрасли. *Инновационные технологии производства и хранения материальных ценностей для государственных нужд*. 2020. № 13. С. 157–163. [Lyndina MI. Ways to improve the food industry. Material assets for government needs: Innovative production and storage. 2020;(13):157–163. (In Russ.)] <https://elibrary.ru/DHQUP>
6. Попов А. М. Анализ и синтез технологий гранулированных концентратов напитков. Кемерово: КемТИПП; 2003. С. 133–148. [Popov AM. Granulated beverage concentrates: Technological analysis and synthesis. Kemerovo: KemTIPP; 2003. pp. 133–148. (In Russ.)]
7. Ксенофонтов В. В. Нейронные сети. Проблемы науки. 2020. № 11. С. 28–29. [Ksenofontov VV. Neural networks. *Problems of science*. 2020;(11):28–29. (In Russ.)] <https://elibrary.ru/DTVJNS>
8. Hassan E, El-Rashidy N, Talaat FM. Mask R-CNN models. *Nile Journal of Communication and Computer Science*. 2022;3(1):17–27. <https://doi.org/10.21608/njccs.2022.280047>
9. Olorunshola OE, Jemitola PO, Ademuwagun A. Comparative study of some deep learning object detection algorithms: R-CNN, fast R-CNN, faster R-CNN, SSD, and YOLO. *Nile Journal of Engineering and Applied Science*. 2023;1(1):70–80. <https://doi.org/10.5455/NJEAS.150264>
10. Prakash SR, Singh PN. Object detection through region proposal based techniques. *Materials Today: Proceedings*. 2021;46(9):3997–4002. <https://doi.org/10.1016/j.matpr.2021.02.533>
11. Zhu L, Lee F, Cai J, Yu H, Chen Q. An improved feature pyramid network for object detection. *Neurocomputing*. 2022;483:127–139. <https://doi.org/10.1016/j.neucom.2022.02.016>
12. Xavier AI, Villavicencio C, Macrohon JJ, Jeng J-H, Hsieh J-G. Object detection via gradient-based mask R-CNN using machine learning algorithms. *Machines*. 2022;10(5):340. <https://doi.org/10.3390/machines10050340>
13. Jabir B, Falih N, Rahmani K. Accuracy and efficiency comparison of object detection open-source models. *International Journal of Online and Biomedical Engineering*. 2021;17(5):165–184. <https://doi.org/10.3991/ijoe.v17i05.21833>
14. Peng H, Yu S. A systematic IOU-related method: Beyond simplified regression for better localization. *IEEE Transactions on Image Processing*. 2021;30:5032–5044. <https://doi.org/10.1109/TIP.2021.3077144>

### Дополнительная информация об авторах / Additional information about the authors

- Шафрай Антон Валерьевич / Anton V. Shafrai ORCID 0000-0003-4512-1933; eLIBRARY SPIN 1840-5242  
Попов Анатолий Михайлович / Anatoliy M. Popov ORCID 0000-0003-0728-7211; eLIBRARY SPIN 6259-7555  
Косинов Виталий Сергеевич / Vitaly S. Kosinov ORCID 0009-0005-5852-9809; eLIBRARY SPIN 8397-7880  
Резниченко Ирина Юрьевна / Irina Yu. Reznichenko ORCID 0000-0002-7486-4704; eLIBRARY SPIN 9847-5357  
Бондарчук Ольга Николаевна / Olga N. Bondarchuk ORCID 0009-0000-0820-5965; eLIBRARY SPIN 8789-9262