

УДК 004.032.26:004.932

Рыжих Владислав Александрович,
Инженер Центрального Кампуса
УНПК «Международный Университет Кыргызстана»
Email: pandorafox@inbox.ru

Рыжих Владислав Александрович,
Борбордук кампустун инженери
ОИӨК «Кыргызстан эл аралык университети»
Email: pandorafox@inbox.ru

Ryzhikh Vladislav Alexandrovich
Central Campus engineer
ERPC "International University of Kyrgyzstan"
Email: pandorafox@inbox.ru

ПРИМЕНЕНИЕ МОДУЛЕЙ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ ПО ЦВЕТОВЫМ ПРИЗНАКАМ

ТҮС ӨЗГӨЧӨЛҮКТӨРҮ БОЮНЧА СҮРӨТТӨРДҮ КЛАССИФИКАЦИЯЛООҮЧҮН ЖАСАЛМА НЕЙРОН ТАРМАГЫНЫН МОДУЛДАРЫН КОЛДОНУУ

APPLICATION OF ARTIFICIAL NEURAL NETWORK MODULES FOR IMAGE CLASSIFICATION BY COLOR FEATURES

Аннотация: В статье рассматривается использование модулей искусственных нейронных сетей для автоматической классификации изображений по их цветовому содержанию. Разработанная модель включает в себя предварительную обработку изображений, построение архитектуры свёрточной нейронной сети и обучение модели на специализированных наборах данных. Результаты демонстрируют высокую точность распознавания цветовых признаков и подтверждают эффективность предложенного подхода.

Ключевые слова: искусственная нейронная сеть, классификация изображений, цветовые признаки, свёрточная нейронная сеть, обучение с учителем.

Аннотациясы: Бул макалада сүрөттөрдү түстүк өзгөчөлүктөрү боюнча автоматтык түрдө классификациялоо үчүн жасалма нейрон тармактарынын модулдарын колдонуу каралат. Иштелип чыккан модель сүрөттөрдү алдын ала иштетүүнү, конволюциялык нейрон тармагынын архитектурасын түзүүнү жана моделди атайын маалымат топтомдоруна үйрөтүүнү камтыйт. Натыйжалар түстүү белгилерди таанууда жогорку тактыкты көрсөтүп, сунушталган ыкманын натыйжалуулугун тастыктайт.

Негизги сөздөр: жасалма нейрон тармак, сүрөттү классификациялоо, түстүү өзгөчөлүктөр, конволюциялык нейрон тармак, көзөмөлдөнгөн окутуу.

Abstract: This article discusses the use of artificial neural network modules for the automatic classification of images based on their color characteristics. The developed model includes preprocessing

of images, construction of a convolutional neural network architecture, and training on specialized datasets. The results demonstrate high accuracy in recognizing color features and confirm the effectiveness of the proposed approach.

Keywords: *artificial neural network, image classification, color features, convolutional neural network, supervised learning.*

Актуальность исследования. Современные методы обработки изображений приобретают всё большее значение в различных областях науки и техники, включая медицинскую диагностику, промышленный контроль качества, Агро промышленность и безопасность. Одним из перспективных направлений является классификация изображений на основе их цветовых признаков, что позволяет эффективно решать задачи автоматической сортировки объектов, обнаружения дефектов и оптимизации технологических процессов. Традиционные алгоритмы обработки изображений на основе ручной настройки цветовых фильтров часто оказываются недостаточно точными или требуют значительных усилий по адаптации под изменяющиеся условия. В этом контексте применение модулей искусственных нейронных сетей, способных самостоятельно выделять и анализировать цветовые особенности изображений, открывает новые возможности для повышения точности и автоматизации процессов. Использование свёрточных нейронных сетей (CNN) позволяет не только повысить качество классификации, но и существенно сократить время анализа за счёт параллельной обработки изображений. В условиях роста объемов данных и необходимости быстрого принятия решений такие преимущества являются критически важными. Таким образом, разработка и применение нейронных сетевых модулей для классификации изображений по цветовым признакам представляют собой актуальную и востребованную задачу, соответствующую современным трендам развития искусственного интеллекта и автоматизации.

Анализ методов и описание полученных результатов

Данное исследование посвящено применению методов компьютерного зрения и машинного обучения для решения задач автоматической классификации изображений объектов (бананов) в реальном времени. Основной целью работы является разработка системы, способной эффективно различать три категории объектов: свежие бананы, гнилые бананы и объекты, не являющиеся бананами.

В рамках поставленной цели решаются следующие задачи: проектирование нейронной сетевой модели, обеспечивающей распознавание объектов заданной категории в различных состояниях (свежие, гнилые); формирование базы данных, включающей изображения бананов в разных состояниях, а также изображения посторонних объектов, с целью повышения точности классификации; оптимизация архитектуры нейронной сети и проведение анализа качества модели для достижения высокой точности распознавания. Предложенная методология основана на интеграции современных методов обработки изображений и глубокого обучения. В процессе исследования применялись: ImageDataGenerator (Keras) для динамической аугментации и нормализации изображений на этапе подготовки данных; YOLO для предварительной фильтрации сцены и вырезания целевых объектов (бананов) из общего изображения; OpenCV (cv2.resize) для приведения размеров изображений к единому формату; TensorFlow/Keras img_to_array для преобразования изображений в формат, пригодный для подачи в нейронную сеть. Особое внимание уделялось созданию структуры нейронной сети, способной поддерживать высокую точность классификации при обработке изображений в реальном времени. Также были применены техники балансировки классов и адаптивного изменения скорости обучения для повышения стабильности модели на практике. Разработанная система направлена на решение задачи автоматической классификации изображений, что позволяет расширить возможности контроля качества на производственных предприятиях и автоматизировать процессы сортировки товаров в условиях увеличивающегося объема данных. Для реализации поставленных задач была разработана комплексная методика, включающая в себя этапы формирования базы данных,

предобработки изображений, выбора и обучения модели нейронной сети, а также её последующего тестирования в различных сценариях.

2.1. Сбор и подготовка данных

Первый этап исследования основан на сборе разнообразного набора изображений. Для этого сформировали собственную базу данных, в который входит 1000 фотографий, разбитых на три основных категории:

- Fresh (Свежие бананы) - 40% выборки;
- Rotten (Гнилые бананы) - 30% выборки;
- Non-fruit (Не бананы) - 30% выборки.

Данное соотношение категорий было выбрано во избежание эффекта переобучения модели и разбалансировки классов на одну категорию и для повышения её способности более эффективно отличать цветовой сектор объектов. Предварительная подготовка данных включает в себя следующие этапы: подгонка изображений к единому размеру 128x128 пикселей, для ускорения обучения и унификации данных; применение методов аугментации данных (повороты, изменения цветового спектра, отражения, яркость, контрастность) для расширения входных данных и повышения устойчивости модели в случаях изменения условий освещения; балансировка изображений, равномерное обучение НС и предотвращение ее смещения в сторону многочисленного класса.

2.2. Детекция объектов с помощью YOLOv5

Для улучшения взаимодействия камеры в реальном времени и во избежание случайных срабатываний, в проект внедрена модель YOLOv5, которая делает предварительный поиск объектов на изображении. Обнаруженный камерой объект во время работы передается в классификационную модель для определения к одному из трех категорий: свежие бананы, гнилые бананы, не бананы. YOLOv5 была выбрана за её высокую скорость работы на CPU и точное определения объектов различных масштабов. Данный подход (YOLO + классификатор) повысил устойчивость к фоновому шуму и человеческому фактору (например, руки в кадре).

1. Архитектура и обучение нейронной сети

Для решения поставленной задачи была построена свёрточная нейронная сеть (CNN), что позволило поднять эффективность задачи компьютерного зрения, особенно в распознавании и классификации различных визуальных паттернов. CNN может извлекать признаки из данного ей изображения путём последовательного применения свёрток и операций субдискретизации (Pooling - уменьшение размера данных).

Входной слой (Input Layer). Все входящие изображения что ей предлагаются унифицируются до размеров 128 x 128 пикселей несмотря на то, что изначально они были разного размера, с тремя цветовыми каналами RGB, что является оптимальным балансом для производительности признаков.

Свёрточный слой (Convolutional Layers). Берутся три последних блока, которые содержат: Свёрточные слои с фильтрами 32, 64 и 128, размер ядра 3x3, который извлекает цветовые и текстовые признаки;

Функция активации ReLU, повышает нелинейность и ускоряет обучение НС;

Слой субдискретизации (MaxPooling) размером 2x2, уменьшает количество размера данных и количество вычислений.

Полностью связанные слои (Full Connected Layers). Два слоя содержат 128 и 64 нейрона предназначенные для обобщения и объединения признаков в один единый вектор. Используется Dropout (с уровнем 0.4) между полностью связанными слоями, который предотвращает переобучение и увеличивает устойчивость модели.

Выходной слой (Output Layer) Содержит 3 нейрона, который соответствует классам «свежие бананы», «гнилые бананы», «не бананы». Для классификации активации применяется функция softmax она позволяет получить вероятность распределения по классам.

В процессе обучения НС были выбраны методы «Функции потерь» и «Оптимизатор Adam». Функция потерь эффективна при решении задач многоклассовой классификации, а оптимизатор Adam ускоряет выполнение задач [Kingma, D.P., Ba, J. 2014]. Цикл обучения НС составляет 150 эпох, а размер мини батча (компромисс между обычным градиентным спуском и стохастическим) 32 изображения. Данные значения позволили получить оптимальное соотношение цикла жизни и обеспечили эффективное использование вычислительных ресурсов системы и стабильной сходимости. Каждое изображение, выделенное YOLOv5, передаётся в классификационную модель на основе сверточной нейронной сети CNN, такой гибридный подход значительно повышает точность системы.

Обучение осуществлялось на заранее подготовленных материалах (база данных), состоящих из 1000 примеров, разделенных на тренировочную (80%) и вариационную (20%) выборки. НС первым делом подгоняет все изображения к одному масштабу (128x18), вторым этапом производит аугментацию данных, а именно генерирует дополнительные варианты изображений путем поворота до 25 градусов, изменяет яркость на 20%, меняет контрастность до 15%, также меняет цветовой баланс. В дальнейшем происходит балансировка классов, автоматическое выравнивание пропорции изображений для предотвращения перекоса модели.

Основным этапом разработки стало использование НС при анализе изображения в реальном времени, используя камеру, подключённую к персональному компьютеру.

Процесс работы модели в реальном времени выглядит следующим образом:

- При помощи библиотеки Open CV была внедрена система видеопотока с камер, которая выполняет захват кадров со скоростью 30 кадров в секунду.
- Каждый кадр сразу проходит предварительную обработку и изменяет размер изображения и происходит нормализация каналов.
- Готовое изображение передаётся в модель, которая сразу выдает результат классификации.
- Готовый результат выводится на экран устройства и отправляет команду на выполнение действия.
- Благодаря быстрой работе CNN и предварительной обработке данных, программа демонстрирует высокую скорость распознавания.

Для оценки устойчивости модели НС был проведен эксперимент в условиях различного освещения, расстояние до объекта от 20 до 50 см, на однородных и неоднородных фонах. НС не только успешно классифицирует изображения на статичных данных, но и обеспечивает очень надежную работу в режиме реального времени. Во время проверки эффективности НС была подготовлена тестовая выборка, состоящая из 1000 изображений, которые не участвовали в обучающей выборке. На тесте модель показала 91,4% общую точность и 88,6% точность распознавания изображения в реальном времени. Среднее время распознавания одного изображения составляет 0,05 секунд. Количество кадров, обработанных за 2,5 минуты, было около 704 кадров.

В процессе использовалась формула вычисления производительности обработки в реальном времени (Bishop, C.M. (2006)) [1]:

$$FPS = \frac{n}{t} \quad [1]$$

где:

- FPS (Frames per second) - производительность обработки изображений,
- n - количество обработки кадров, - t - общее время обработки в секунду. В нашем случае:

$$FPS = \frac{704}{60} = 11,7 \text{ кадров/секунду}$$

3.1. Математическое описание архитектуры нейросети и процесса классификации 1)

Общая формула вывода слоя CNN (Chollet, 2017) [2]:

$$Y^{(l)} = f\left(\sum_{i=1}^M X_i^{(l-1)} \otimes W_i^{(l)} + b^{(l)}\right) \quad [2]$$

где:

- $Y^{(l)}$ - $(l-1)$ выход текущего сверточного слоя;
- X_i - выходные карты признаков с предыдущего слоя;
- $W_i^{(l)}$ - веса (ядро свёртки) текущего слоя;
- $b^{(l)}$ - вектор смещения (bias);
- $f(*)$ - нелинейная функция активации (ReLU в нашем случае);
- \otimes - операция свертки.

Каждый нейросетевой слой постепенно обобщает и извлекает признаки изображения, уменьшая размерность и увеличивая абстрактность представлений, что позволяет эффективно классифицировать изображения на классы: свежие бананы, гнилые бананы, не бананы.

2) Выходной слой определяется формулой функции softmax (Bishop, 2006) [3]:

$$P(y = i|x) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad [3]$$

где:

- $P(y = j|x)$ - вероятность того, что входное изображение x относится к классу j ;
- Z_j - выход нейрона, соответствующего классу j ;
- K - общее количество классов.

Каждому классу присваивается вероятность принадлежности, и финальная классификация осуществляется по максимальному значению вероятности (Bishop, С.М. (2006)) [4]:

$$\text{класс изображения} = \arg \min_j (P(y=j|x)) \quad [4]$$

Формула позволяет присваивать изображениям классы и количественно оценивать уверенность в принятом решении.

4. Практическое применение Графики

точности и потерь:

- график продемонстрировал рост точности, что показывает правильность выбранной архитектуры, которая достигается максимума на 50-й эпохе (91.4%) (Рисунок 1).

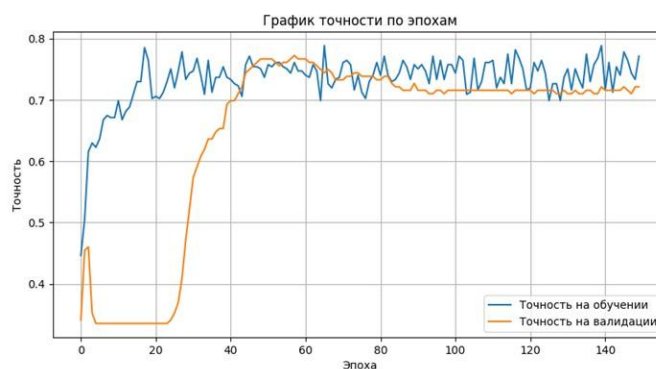


Рисунок 1 - График точности нейросети по эпохам обучения (Источник: составлено автором)

- график показывает уверенное снижение ошибок с начала каждого цикла эпохи, что говорит нам о корректной работе оптимизатора (Рисунок 2).

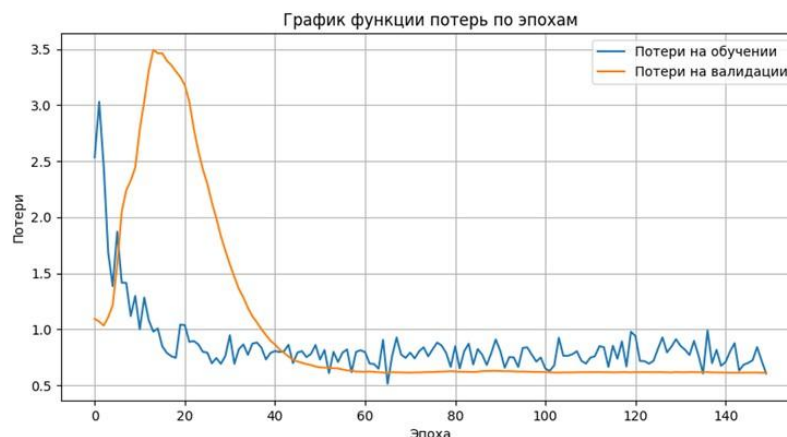


Рисунок 2 - График функции потерь нейросети по эпохам обучения (*Источник: составлено автором*)

В процессе обучения использовалась функция потерь (категориальная кроссэнтропия), вычисляемая по формуле (Goodfellow et al., 2016) [5]:

$$L(y, \hat{y}) = - \sum_{i=1}^n y_i * \log(\hat{y}_i) \quad (5)$$

где:

- y_i - истинное значение класса (0 или 1),
- \hat{y}_i - предсказанное моделью значение вероятности принадлежности класса, - n - общее число классов.

Минимизация этой функции потерь позволяет нейронной сети эффективно корректировать веса и обучаться классифицировать изображения с высокой точностью.

Заключение

В результате исследования создана программа, которая способна быстро и точно отличать бананы от других объектов, не являющихся фруктами, а также классифицировать их по трем признакам: свежие, гнилые и не бананы. В разработке использовались современные свёрточные НС для обеспечения стабильной работы в условиях разного освещения и изменении углов обзора.

Модель показала высокий уровень устойчивости и точности при классификации изображений по цветовым признакам. В условиях слабого освещения с перекрытием объекта, на изображении проявились ошибки. Это значит что в дальнейшем надо усовершенствовать алгоритмы обработки и повысить точность модели к изменениям условий слабого освещения. С учетом полученных данных можно сформулировать следующие заключения: ошибки модели допускаются при классификации изображений, относящихся к классу не бананы, особенно когда объекты имели форму и цвет, схожие с бананами; модель различает состояние фруктов (свежий и гнилой) тем самым демонстрирует минимальное количество ошибок в этих классах.

Практическая значимость разработанной НС заключается в возможности интегрировать её в автоматизированные системы сортировки фруктов, в контроле качества в агропромышленности и розничной торговле. В дальнейшем планируется улучшение базы данных и создание мобильного приложения, для расширения области ее применения. Это позволит использовать данную разработку как в промышленном, так и в повседневном бытовом использовании.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ:

1. Bishop, C.M. (2006). Pattern Recognition and Machine Learning. Springer, -738 p.
2. Chollet, F. (2017). Deep Learning with Python. Manning Publications, -384 p.
3. Goodfellow, I., Bengio, Y. and Courville, A. (2016). Deep Learning. -MIT Press, 8-00 p.

4. Kingma, D.P. and Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.
5. Raschka, S. (2015). Python Machine Learning: Unlock deeper insights into Machine Learning with this vital guide to cutting-edge predictive analytics. -Packt Publishing, -454 p.