

А.Е. Аязбай¹, Г.К. Сергазин², С.К. Оразалиева¹, Б.С. Байкенов¹

¹Алматинский университет энергетики и связи имени Г. Даукеева, Алматы, Казахстан

²Академия логистики и транспорта, Алматы, Казахстан

*E-mail: work_abu@hotmail.com

ОБУЧЕНИЕ МОДЕЛИ YOLO ДЛЯ РАСПОЗНАВАНИЯ ЛЕДЕНЦОВ HALLS С ПОМОЩЬЮ MAIXDUINO

Аннотация. В данном исследовании рассматривается использование платы разработки MaixDuino для распознавания трех различных сортов леденцов (конфет). Этап обучения проводился на платформе MaixHub, представляющей собой комплексную экосистему, объединяющую сервисы ИИ-моделей и взаимодействие с сообществом. Работая в тандеме с моделью YOLOv2, MaixDuino прошла пять итераций обучения, постепенно улучшая свои характеристики.

Наибольшее внимания заслуживает пятая, заключительная тренировка, показавшая оптимальную точность. Исследование показало, что модель демонстрирует отличные результаты в реальных условиях, особенно при хорошем освещении и отсутствии бликов. Предполагается, что данная система обнаружения леденцов будет интегрирована в процесс сортировки с использованием комбинации ленточного конвейера и дельта-манипулятора.

Данное исследование вносит существенный вклад в развитие технологий обнаружения леденцов, демонстрируя эффективность платы разработки MaixDuino в сочетании с моделью YOLOv2. Результаты исследования подчеркивают возможность применения данной системы в промышленных условиях, где надежная и эффективная сортировка конфет имеет первостепенное значение.

Ключевые слова. MaixDuino, YOLOv2, распознавание, машинное зрение, обработка изображений, промышленная сортировка.

Введение.

В области автоматизированного обнаружения и сортировки объектов интеграция технологий машинного зрения и искусственного интеллекта (ИИ) приобретает ключевое значение для повышения эффективности и точности.

Исследование [1] было посвящено разработке и оптимизации недорогого автоматического датчика цвета RGB для сортировки фруктов. Авторы использовали многоцелевую систему, состоящую из аппаратных компонентов, таких как рама машины, конвейерная лента и датчики цвета RGB, а также программных компонентов, включая Arduino IDE и C++.

В работе [2] рассматривались технологии распознавания внешнего вида, сортировки и перегруппировки аккумуляторов. В исследовании использовалось машинное зрение для фильтрации внешнего вида и аналитический иерархический процесс, и анализ отношений Грея для классификации на основе сценариев применения.

В работе [3] авторы предложили систему на основе машинного зрения с использованием аппаратной архитектуры FPGA для сортировки цитрусовых. Система продемонстрировала высокую точность при сегментации, классификации по цвету и размеру фруктов, достигнув 97%, 94% и 90% точности соответственно.

Для решения проблемы качества семян хлопчатника в хлопковой промышленности в работе [4] предложен метод обнаружения повреждений семян хлопчатника по изображениям, основанный на усовершенствованном алгоритме YOLOv5. Алгоритм

позволил достичь высокой точности при идентификации повреждений по внешнему виду, превосходя традиционные методы по точности и скорости обработки.

В работе [5] была представлена алгоритмическая архитектура для обнаружения мусора и трехмерной пространственной локализации. Предложенный подход, основанный на передовых методах компьютерного зрения и машинного обучения, позволяет в реальном времени выполнять сегментацию объектов, точное распознавание мусора и пространственную локализацию с максимальной скоростью 150 кадров в секунду. Результаты экспериментов продемонстрировали эффективность предложенного метода для преобразования процедур утилизации отходов.

В данном исследовании решается задача идентификации и классификации трех различных типов конфет с помощью платы разработки MaixDuino - недорогого и компактного решения для машинного зрения. Основной целью является применение модели YOLOv2 (You Only Look Once), хорошо зарекомендовавшей себя архитектуры конволюционной нейронной сети (CNN), для точного обнаружения конфет в системе сортировки в реальном времени.

Хотя YOLOv2 можно считать устаревшей моделью в быстро развивающейся области компьютерного зрения, ее возможности полностью соответствуют нашим конкретным задачам и хорошо поддерживаются платформой MaixDuino. Плата разработки MaixDuino, известная своей доступностью и компактностью, оказывается идеальным аппаратным решением для приложений ИИ. Такая интеграция позволяет выполнять алгоритмы ИИ на устройствах с ограниченными ресурсами, что делает ее особенно подходящей для сценариев обработки данных в реальном времени.

В работе [7] авторы использовали плату MaixBit для визуальной навигации дронов. Плата использовалась для обнаружения подсолнечника. В плате MaixBit используется та же система на чипе, что и в MaixDuino – Kendryte K210.

В работе [8] авторы предложили систему обнаружения и тушения пожара, которая была разработана с использованием методов обработки изображений и платы разработки MaixDuino и направлена на повышение пожарной безопасности путем анализа изображений в реальном времени.

Разработанная в работе [9] система для обнаружения масок, использующее камеру для ввода фото- и видеoinформации, подключенную к микропроцессору Sipeed Maix Bit, предназначено для наблюдения в режиме реального времени за людьми, не использующими маски в местах общественного пользования. Исследование завершается успешным экспериментом, демонстрирующим высокую точность обнаружения лиц, не использующих маски в общественных местах, что позволяет внести ценный вклад в разработку и реализацию основанного на правилах детектора масок для профилактики COVID-19.

Интеграция MaixDuino и YOLOv2 отражает стремление сделать ИИ доступным, особенно в сценариях, где экономическая эффективность и компактность имеют первостепенное значение. В данном исследовании используется платформа MaixHub, предоставляемая компанией Sipeed, что способствует легкому доступу к сервисам моделей ИИ и развитию сотрудничества между сообществами.

Значение данного исследования выходит за рамки технического прогресса, достигнутого в результате внедрения YOLOv2, и имеет практическое значение для промышленности. Возможности YOLOv2 в реальном времени в сочетании с ресурсосберегающим исполнением на платформе MaixDuino позволяют повысить эффективность производства, снизить количество ошибок и обеспечить стабильное качество продукции. Кроме того, адаптивность обученных моделей к различным условиям окружающей среды и сценариям освещения имеет решающее значение для их успешного применения в различных производственных условиях.

При детальном рассмотрении процесса обучения и результатов в последующих разделах становится очевидным, что итерационный характер тренировок способствует совершенствованию модели. Этот итеративный процесс приводит к оптимальной работе в реальных условиях, демонстрируя синергию между хорошо зарекомендовавшей себя моделью YOLOv2 и доступным, недорогим аппаратным решением MaixDuino. Результаты данного исследования не только продвигают современный уровень автоматизированного обнаружения конфет, но и демонстрируют более широкий потенциал ИИ в системах распознавания и сортировки объектов в различных отраслях промышленности.

Материалы и методы.

Методология обучения модели YOLOv2 для платы разработки MaixDuino, предназначенной для обнаружения конфет, включает в себя системный подход, объединяющий ключевые параметры и конфигурации для достижения оптимальной производительности. В процессе обучения учитываются размеры наборов данных, количество эпох обучения и конкретные конфигурации, применяемые при обучении модели.

В качестве основной среды для обучения модели использовалась платформа MaixHub, представленная компанией Sipeed. Благодаря использованию сервиса онлайн-обучения на MaixHub процесс обучения был облегчен, что обеспечило доступность для пользователей с разным уровнем владения искусственным интеллектом.

Набор обучающих данных состоит из обширной коллекции изображений, отражающих различные визуальные характеристики трех разных видов конфет. Изображения были получены как с камер MaixDuino, так и с камер смартфонов, что позволило получить разнообразную картину реальных ситуаций.

Набор данных итеративно расширялся в ходе нескольких сессий обучения. В первых экспериментах использовался набор из 276 изображений, в последующих сессиях он увеличивался до 307, 471 и 681 изображения. Такое постепенное расширение было направлено на повышение адаптивности модели к разным условиям. Пример аннотированного набора данных представлен на рисунке 1.

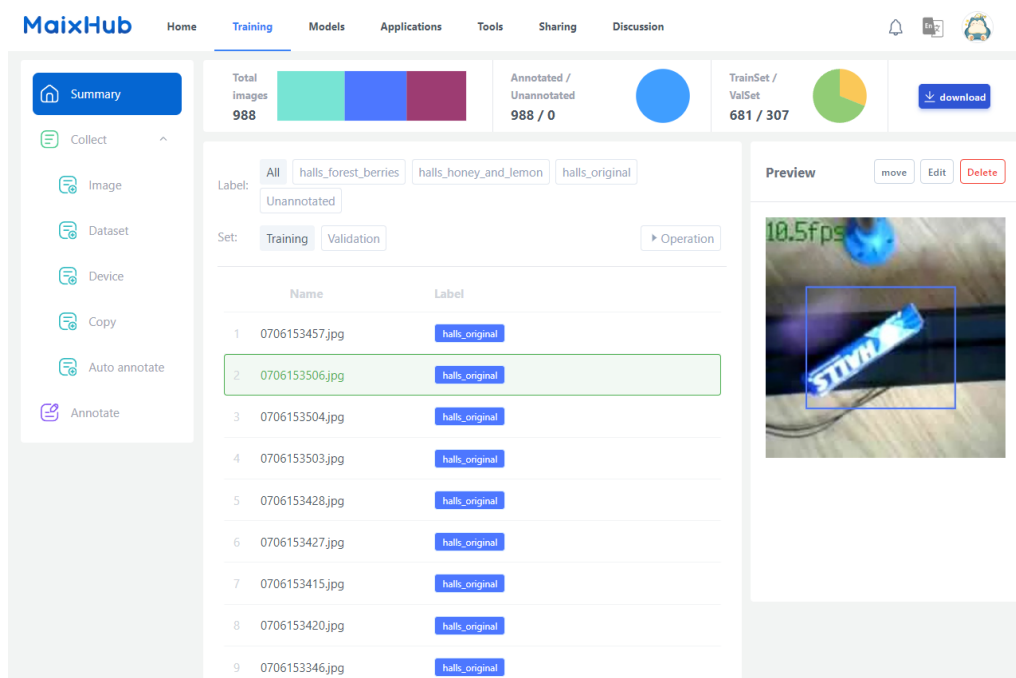


Рисунок 1 – Загруженные и аннотированные изображения наборов данных

Сеансы обучения проводились в течение заданного количества эпох, представляющих собой полные проходы по всему обучающему набору данных. Динамика производительности модели наблюдалась на протяжении 100 и 200 эпох в разных сеансах обучения. Это позволило определить оптимальную точку сходимости модели при различных размерах и конфигурациях набора данных.

Минимальный размер ограничительных рамок, определяющих наименьшую различимую область конфеты, систематически изменялся в процессе обучения. Вариации включали в себя настройки с минимальными размерами рамок 5 и 10 пикселей, что отражало желаемую гранулярность обнаружения в процессе обучения модели.

Эффективность модели оценивалась по ключевым метрикам, включая точность, потери и точность валидационного набора (accuracy, loss, validation set accuracy). Точность, представляющая собой среднюю точность, позволяет оценить сходимость модели в процессе обучения. Точность на валидационном множестве, являющаяся важнейшей метрикой, дает представление о прогностических возможностях модели на ранее не изученных данных, что способствует комплексной оценке эффективности ее обобщения.

Описанная выше методология создает структурированную основу для обучения модели YOLOv2 на плате MaixDuino. Последующие разделы раскрывают результаты этих тренировок, проливая свет на эволюцию эффективности модели при различных объемах и конфигурациях наборов данных.

Результаты.

В настоящем документе представлены результаты тренировок модели YOLOv2 для платы MaixDuino, в которых подробно рассматривается динамика изменения показаний модели при различных объемах наборов данных и конфигурациях.

В первой сессии обучения на наборе данных из 276 изображений модель продемонстрировала многообещающую сходимость. Точность валидационного набора достигла 0,84848 на эпохе 60, потери составили 0,00374, а общая точность – 0,30926. Примечательно, что эта сессия позволила получить фундаментальное представление о поведении модели и заложить основу для последующих улучшений. Результаты обучения представлены на рисунке 2.

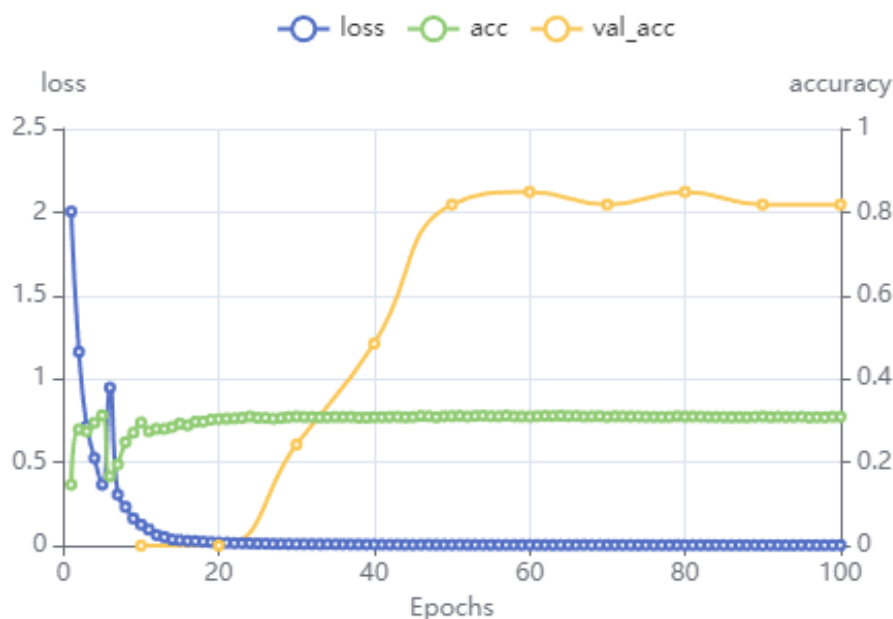


Рисунок 2 – Результаты первой сессии обучения

Вторая обучающая сессия, сохранив размер набора данных в 276 изображений, но увеличив количество эпох до 200 и уменьшив минимальный размер рамки до 5 пикселей, продемонстрировала существенное улучшение. Точность валидационного набора увеличилась до 0,93939 на эпохе 80, при этом потери составили 0,00216, а точность – 0,19397. Эти корректировки заметно повысили прогностические возможности модели. Результаты обучения представлены на рисунке 3.



Рисунок 3 – Результаты второй сессии обучения

На расширенном наборе данных, состоящем из 307 изображений и 200 эпох обучения, в третьей сессии были получены убедительные результаты. Точность валидационного набора достигла идеального значения 1,0 на эпохе 120, потери составили 0,00113, а точность - 0,20896. Увеличение объема набора данных способствовало повышению универсализации модели. Результаты обучения представлены на рисунке 4.



Рисунок 4 – Результаты третьей сессии обучения

Четвертая сессия обучения включала 471 учебное изображение и предусматривала минимальный размер блока в 10 пикселей в течение 200 эпох. Примечательно, что модель достигла идеальной точности валидационного набора 1,0 при точности 0,24744 и потере

0,00735 на эпохе 60. Расширение набора данных позволило уточнить дискриминационные возможности модели. Результаты обучения представлены на рисунке 5.



Рисунок 5 – Результаты четвертой сессии обучения

Пятая, заключительная обучающая сессия, в которой использовался набор данных из 681 изображения, продемонстрировала наиболее заметное улучшение производительности. Точность валидационного набора достигла идеального значения 1,0 на эпохе 100, при этом точность составила 0,38923, а потери - 0,0036. Эта сессия с расширенным набором данных оказалась наиболее эффективной в реальных условиях. Результаты обучения представлены на рисунке 6.



Рисунок 6 – Результаты пятой сессии обучения

Эти совокупные результаты свидетельствуют об итерационном совершенствовании модели YOLOv2 в ходе тренировочных сессий, показывая рост эффективности при расширении набора данных и выборе оптимальных параметров. Пятая обучающая сессия, характеризующаяся самым большим набором данных, стала наиболее успешной, продемонстрировав оптимальную производительность в реальных сценариях. В последующих разделах будут рассмотрены последствия этих результатов обучения в контексте предполагаемого применения для распознавания и сортировки конфет.

Примеры распознавания представлены на рисунке 7.

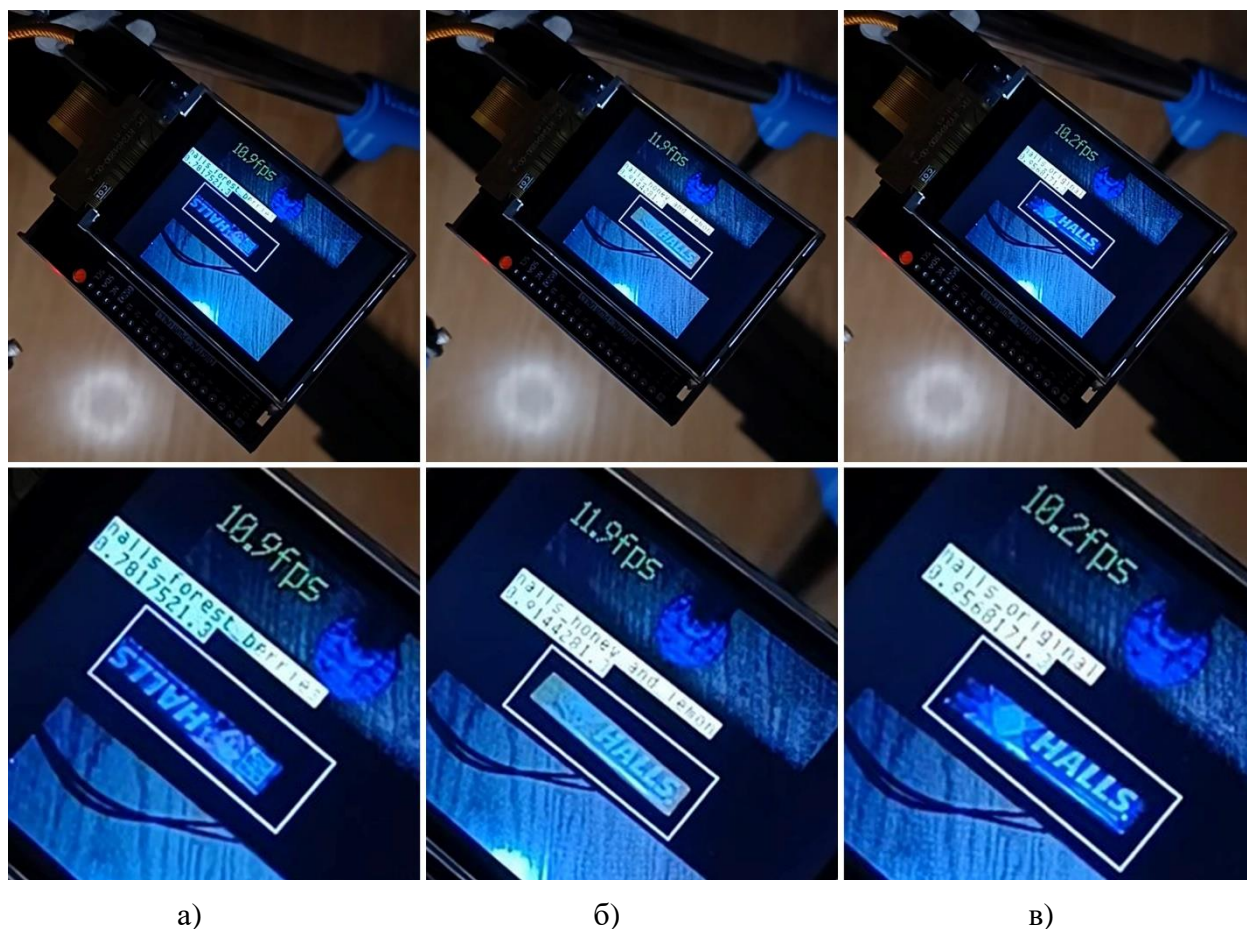


Рисунок 7 – Результат обнаружения трех различных вкусов конфет Halls (а – лесные ягоды, б – мед и лимон, в – оригинал)

Обсуждение.

Наблюдаемые тенденции в производительности модели в разных сессиях обучения подчеркивают чувствительность модели YOLOv2 к изменениям в размерах набора данных и конфигурациях обучения. Примечательно, что итеративное расширение набора данных неизменно способствовало повышению точности валидационного набора, что свидетельствует об улучшении возможностей универсализации.

При постепенном увеличении набора данных с 276 до 681 изображения наблюдалось пропорциональное повышение точности модели. Наблюдаемая корреляция между размером набора данных и производительностью модели согласуется с основополагающими принципами глубокого обучения, подчеркивающими важность разнообразных и обширных обучающих данных для надежного обобщения модели.

В ходе обучения выявлено сложное взаимодействие между количеством эпох обучения и минимальным размером ограничивающего поля. Оптимальная сходимость модели часто достигалась при увеличении числа эпох и уменьшении минимального размера блока, что подчеркивает необходимость тонкой настройки параметров для достижения оптимальной производительности.

Несмотря на то, что пятая тренировка показала наивысшую точность, необходимо отметить контролируемые условия, в которых были получены эти результаты. Применимость в реальных условиях может зависеть от таких факторов, как освещение, блики и вариации в подаче конфет. Для подтверждения надежности модели в практических сценариях использования необходимо провести дальнейшие исследования в различных условиях внешней среды.

Обученная модель YOLOv2, особенно после пятой тренировки, может быть интегрирована в сортировочную систему с ленточным конвейером и дельта-манипулятором. Повышенная точность распознавания типов конфет в оптимальных условиях делает модель перспективной для оптимизации и повышения эффективности процессов сортировки конечных изделий.

Обсуждаются также проблемы, присущие модели, в том числе ее чувствительность к факторам окружающей среды и потенциальные ограничения в реальных условиях. Дальнейшие направления работы должны включать в себя доработку модели с целью повышения ее надежности и адаптации к различным условиям эксплуатации.

Сравнительный анализ с современными моделями обнаружения и сортировки конфет позволил бы более точно оценить эффективность модели YOLOv2 на платформе MaixDuino. В будущих исследованиях можно изучить возможность интеграции более современных архитектур и методологий, чтобы убедиться в повышении точности и эффективности.

В заключение следует отметить, что в ходе обсуждения были всесторонне проанализированы результаты обучения, выявлены факторы, влияющие на производительность модели, потенциальные проблемы и направления дальнейших исследований. Это понимание нюансов позволяет расширить дискуссию об использовании YOLOv2 и MaixDuino в области автоматического обнаружения и сортировки конфет.

Заключение.

Итерационные тренировки, отмеченные изменениями в размерах наборов данных и конфигурациях обучения, продемонстрировали заметное улучшение работы модели YOLOv2. Прогресс от первой сессии к пятой показал постепенное увеличение точности валидационного набора, подтвердив способность модели с большей точностью определять тип конфет.

Важным результатом является выраженное влияние размера набора данных на точность модели. Положительная корреляция между увеличением набора данных и улучшением обобщения модели подчеркивает важность разнообразных и обширных обучающих данных для достижения надежных возможностей обнаружения объектов.

Обученная модель YOLOv2, особенно после пятого сеанса обучения, является перспективным кандидатом для интеграции в системы сортировки конечных изделий. Продemonстрированная точность распознавания конфет в оптимальных условиях делает ее ценным инструментом для повышения эффективности и точности процессов сортировки в промышленных условиях.

При этом необходимо учитывать проблемы, связанные с факторами окружающей среды и потенциальными ограничениями в реальных сценариях. Такие аспекты, как перепады освещенности и блики, могут повлиять на устойчивость модели, что требует дальнейшего изучения и доработки для обеспечения всесторонней применимости.

Результаты данного исследования закладывают основу для будущих направлений исследований. Изучение адаптации модели к различным условиям эксплуатации, сравнение с современными моделями и дальнейшее совершенствование с целью повышения надежности представляют собой направления для дальнейших исследований.

Данная работа вносит вклад в более широкую дискуссию о применении ИИ в системах обнаружения и сортировки объектов, подчеркивая потенциал модели YOLOv2 и доступность платформы MaixDuino. Сочетание устоявшейся модели и экономически эффективного, компактного аппаратного решения открывает перспективы для демократизации возможностей ИИ в различных отраслях.

В заключение следует отметить, что совместная интеграция модели YOLOv2 с платой разработки MaixDuino представляет собой значительный шаг на пути к созданию эффективных и доступных приложений ИИ в области обнаружения и сортировки конфет. Полученные результаты не только подчеркивают адаптивность модели, но и заставляют задуматься о ее интеграции в реальные промышленные объекты. Проведенные исследования закладывают основу для дальнейшего прогресса в области пересечения краевого ИИ, машинного зрения и автоматизированных систем сортировки.

ЛИТЕРАТУРА

- [1] Elwakeel A.E., Mazrou Y.S.A., Tantawy A.A., Okasha A.M., Elmetwalli, A.H., Elsayed S., Makhlof A.H. Designing, Optimizing, and Validating a Low-Cost, Multi-Purpose, Automatic System-Based RGB Color Sensor for Sorting Fruits. *Agriculture* 2023, 13, 1824. <https://doi.org/10.3390/agriculture13091824>
- [2] Liu F., Chen J., Qin D., Wang T. Research on Appearance Detection, Sorting, and Regrouping Technology of Retired Batteries for Electric Vehicles. *Sustainability* 2023, 15, 15523. <https://doi.org/10.3390/su152115523>
- [3] Nuño-Maganda M.A., Dávila-Rodríguez I.A., Hernández-Mier Y., Barrón-Zambrano J.H., Elizondo-Leal J.C., Díaz-Manriquez A., Polanco-Martagón S. Real-Time Embedded Vision System for Online Monitoring and Sorting of Citrus Fruits. *Electronics* 2023, 12, 3891. <https://doi.org/10.3390/electronics12183891>
- [4] Liu Z., Wang L., Liu Z., Wang X., Hu C., Xing, J. Detection of Cotton Seed Damage Based on Improved YOLOv5. *Processes* 2023, 11, 2682. <https://doi.org/10.3390/pr11092682>
- [5] Lv Z., Chen T., Cai Z., Chen Z. Machine Learning-Based Garbage Detection and 3D Spatial Localization for Intelligent Robotic Grasp. *Appl. Sci.* 2023, 13, 10018. <https://doi.org/10.3390/app131810018>
- [6] Hulens D., Van Ranst W., Cao Y., Goedemé T. Autonomous Visual Navigation for a Flower Pollination Drone. *Machines* 2022, 10, 364. <https://doi.org/10.3390/machines10050364>
- [7] Vignesh S., Vishnupriya G., Swetha P., Elakkiya, B. Fire detection and extinguisher system using image processing. *International Research Journal of Engineering and Technology* 2023, 10, 6.
- [8] Rafika A. S., Warnar H. L. H. S., Saputra M. A. An application of mask detector for prevent Covid-19 in public services area. In *Journal of Physics: Conference Series*, November 2020, (Vol. 1641, No. 1, p. 012063). IOP Publishing.

Абу-Алим Аязбай, магистр, Ф. Дәукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан, work_abu@hotmail.com

Гани Сергазин, PhD, қауымдастырылған профессор, Логистика және көлік Академиясы, Алматы, Қазақстан, gani_b@mail.ru

Оразалиева Сандугаш, PhD, Ф. Дәукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан, s.orazalieva@aes.kz

Бахытжан Байкенов, т.ғ.к., доцент, Ф. Дәукеев атындағы Алматы энергетика және байланыс университеті, Алматы, Қазақстан, b.baikenov@aes.kz

МАIХDUINO-МЕН HALLS КӘМПИТТЕРІН АНЫҚТАУ ҮШІН YOLO МОДЕЛІН ОҚЫТУ

Аңдатпа. Бұл зерттеу кәмпиттің үш түрін анықтау үшін MaixDuino әзірлеу тақтасын пайдалануды зерттейді. Жаттығу кезеңі MaixHub платформасында өткізілді, ол

AI моделінің қызметтері мен қауымдастықтың өзара әрекеттесуін біріктіретін кешенді экожүйе болып табылады. YOLOv2 үлгісімен тандемде жұмыс істей отырып, MaixDuino бес оқу итерациясынан өтіп, жұмысын біртіндеп жақсартта түсті.

Оңтайлы дәлдікті көрсеткен бесінші және соңғы жағтығу сессиясы ерекше назар аударуға лайық. Зерттеу көрсеткендей, модель нақты әлем жағдайында, әсіресе жақсы жарықтандыруда және жарқыраусыз жағдайларда тамаша нәтижелер көрсетеді. Бұл кәмпиттерді анықтау жүйесі конвейер лентасы мен дельта манипулятордың комбинациясын пайдалана отырып, сұрыптау процесіне біріктіріледі деп болжануда.

Бұл зерттеу MaixDuino әзірлеу тақтасының YOLOv2 үлгісімен үйлесетін тиімділігін көрсету арқылы кәмпиттерді анықтау технологияларының дамуына елеулі үлес қосады. Зерттеу нәтижелері кәмпиттерді сенімді және тиімді сұрыптау маңызды болып табылатын өнеркәсіптік жағдайларда бұл жүйенің орындылығын көрсетеді.

Түйінді сөздер. MaixDuino, YOLOv2, тану, машиналық көру, кескінді өңдеу, өнеркәсіптік сұрыптау.

Abu-Alim Ayazbay, master, Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, work_abu@hotmail.com

Gani Sergazin, PhD, associate professor, Academy of logistics and transport, Almaty, Kazakhstan, gani_b@mail.ru

Sandugash Orazalieva, PhD, Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, s.orazalieva@aes.kz

Bakhytzhан Baikenov, candidate of technical sciences, docent, Almaty University of Power Engineering and Telecommunications named after G. Daukeyev, Almaty, Kazakhstan, b.baikenov@aes.kz

TRAINING YOLO MODEL FOR HALLS CANDIES DETECTION WITH MAIXDUINO

Abstract. This research delves into the utilization of the MaixDuino development board for the detection of three distinct varieties of halls candies. The training phase was carried out on the MaixHub platform, a comprehensive ecosystem integrating AI model services and community interaction. Operating in tandem with the YOLOv2 model, the MaixDuino underwent five training iterations, progressively refining its performance.

The fifth and final training session emerged as the most noteworthy, exhibiting optimal accuracy. The study underscores the model's remarkable real-world performance, particularly in well-lit conditions devoid of glares. The envisioned application involves the integration of this candy detection system into sorting processes, utilizing a combination of a belt conveyor and a delta manipulator.

This research significantly contributes to the landscape of candy detection technologies, showcasing the efficacy of the MaixDuino development board coupled with the YOLOv2 model. The study's findings emphasize the potential applications of this system in industrial settings, where reliable and efficient candy sorting is paramount.

Keywords. MaixDuino, YOLOv2, detection, machine vision, image processing, industrial sorting.
