

Використання комп'ютерного зору для візуальної перевірки якості груш

М. Ільчук, аспірант; А. Стадник, аспірант,

Національний університет «Львівська політехніка», Україна; e-mail: Rexery190@gmail.com

Анотація

Візуальний контроль залишається одним з найважливіших та найшвидших методів неінвазивної перевірки продукції та її класифікації за рівнем якості при виробництві та розподіленні для споживача. Зазвичай цим займаються кваліфіковані оператори, завдання яких швидко прийняти рішення та встановити факт наявності візуального дефекту. Особливо це актуально в харчовій промисловості. Даний метод має очевидну проблему, а саме – людський фактор. Оскільки не можна не враховувати фактор втоми оператора, його суб'єктивність при виборі допустимого рівня відхилення продукту. Для вирішення цього завдання у даній роботі пропонується використати методи комп'ютерного зору при визначенні класу якості, а конкретизуючи, при розподілу груш за класом екстра та поганої якості. Наводиться короткий опис основних етапів обробки зображення для звернення уваги на етап сегментизації, як можливий шлях для підвищення точності при прийнятті рішень. Представлена модель основними характеристиками якої є візуальні ознаки, такі як: колір, форма, наявність стебла та інші. Яким на етапі навчання відновно присвоюються уже числові характеристики. Та за рахунок запропонованих різних підходів при обробці зображення досягається високий рівень точності, який виражається у процентному співвідношенні вірності прийняття рішень та варіюється у діапазоні від 90 до 96%. Тому отримані результати у даній роботі дають можливість автоматизувати процес візуальної перевірки та водночас досягти високої точності прийняття рішень, з перспективою підвищення швидкості та якості реалізації продукції для споживача.

Ключові слова

Контроль якості, Комп'ютерний зір, Автоматизація

1. Вступ

Завдання візуального огляду для розпізнавання об'єктів та оцінки його якості є одним з найважливіших процесів в обробній та харчовій галузях промисловості. З огляду на вимоги замовника, є обов'язковим забезпечення саме високої якості продукту. Питання перевірки об'єктів з метою виявлення дефектів, таких як колір, подряпин, тріщин або перевірки поверхонь на належне покриття, пов'язане з візуальним контролем якості [1]. Однією з альтернатив, прийнятих багатьма галузями для збереження конкурентоспроможності, є сприяння бережливому виробництву, в якому ці методи можуть працювати синергетично для створення спрощеної високоякісної системи, яка виробляє готову продукцію за темпами попиту споживачів з невеликими відходами або без них. На жаль, критикуючи наявні дані, які свідчать про те, що декілька організаційних факторів можуть унеможливити або перешкодити впровадженню методів бережливого виробництва серед виробничих підприємств. Інша альтернатива — надати комп'ютеру можливість автоматичного огляду та розпізнавання об'єктів. На нашу думку використання спеціалізованого програмного забезпечення разом з іншими механізмами, такими як камери, сенсори та залучення висококваліфікованих фахівців дозволяють отримати потужний інструмент для автоматичної швидкої перевірки якості продукції. Така автоматизація надає можливість здешевити та пришвидшити виробництва при високому рівні точності встановлення якості продукту. Саме дана технологія може відігравати важливу роль при перевірці фруктів. Оскільки тут важливою умовою є дотримання неінвазивного методу, а також враховуючи факт, що за візуальним станом можна практично точно вказати його якість.

Комп'ютерний зір (Computer Vision, CV) [2] займається моделюванням та відтворенням людського бачення за допомогою комп'ютерного програмного та апаратного забезпечення. Це також основна теорія для створення штучних систем для вилучення інформації із зображень. Оскільки за аналогією з людиною, яка більше 70% інформації навколишнього світу сприймає саме через зір, дана теорія дозволяє досягти значних успіхів. Головне завдання – правильно інтерпретувати інформацію з зображення яке ми отримуємо.

2. Недоліки

Незважаючи на численні дослідження, розроблені в області комп'ютерного зору, досі не існує стандартизованого методу, який можна було б запропонувати для оцінки якості різних типів об'єктів. Особливі характеристики об'єкта вимагають налаштування системи комп'ютерного зору; це передбачає вичерпний процес дослідження, а не лише придбання дорогого обладнання. Крім того, для досягнення кращої продуктивності системи є обов'язковим отримання якомога ліпших вхідних даних.

3. Мета роботи

Метою дослідження є розроблення автоматизованої моделі візуальної перевірки якості груш, при досягненні більше 90% правильних висновків під час прийняття рішення про стан продукції.

4. Комп'ютерний зір при обробці цифрових зображень. Сучасний підхід.

Задача полягає в представленні короткого опису основних етапів роботи систем компютерного зору з критичним аналізом на якому етапі та за рахунок чого ми можемо досягти поставленої точності моделі та навівши практичні досліди обрати оптимальний варіант.

Виявлення об'єктів. Більшість сучасних підходів виявлення об'єктів виконують регресію обмежувальної рамки та класифікацію категорій на емпіричних апріорних об'єктах, таких як пропозиції, прив'язки, точки [3]. Нещодавно було запропоновано виявляти об'єкти за допомогою фіксованого набору запитів, які можна вивчати. Відтоді парадигма виявлення на основі запитів привернула велику увагу та надихнула на серію робіт. У багатьох роботах це підштовхує подальший розвиток виявлення об'єктів. Що продемонстровано на рисунку 1.

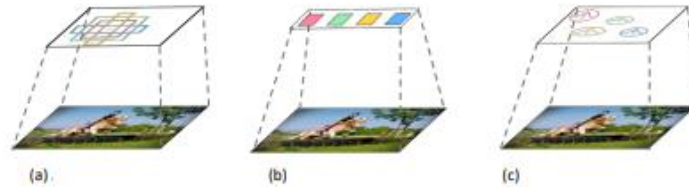


Рисунок 1. Порівняння різних парадигм виявлення об'єктів.
(a) Виявлення з попередніх емпіричних об'єктів; (b) Виявлення із запитів, що навчаються; (c) Виявлення на основі випадкових боксів. [3].

Впровадження систем комп'ютерного зору значно залежить від області їхнього застосування [4]. Реалізація систем комп'ютерного зору також залежить від того, є її функціональність наперед визначеною чи деякі її частини можуть бути вивчені і вдосконалені в процесі роботи. Однак існують функції, типові для багатьох систем комп'ютерного зору. З врахуванням специфіки та заданої мети на кожному етапі було введено корективи, які описуються нижче, з метою досягнення зазначеної точності при забезпеченні оптимальної швидкості роботи моделі.

Варто також згадати про Масковані моделі світу (MWM) [5], алгоритм на основі візуальної моделі, який роз'єднує навчання візуального представлення та навчання динаміки. Ключова ідея MWM полягає в тому, щоб навчити автокодувальник, який реконструює візуальні спостереження за допомогою згорткового маскування ознак і моделі прихованої динаміки поверх автокодера. Завдяки введенню ранніх згорткових шарів і маскуванню згорткових функцій замість піксельних плям наш підхід дозволяє моделі світу охоплювати дрібні візуальні деталі зі складних візуальних спостережень. Крім того, для того, щоб дізнатися інформацію, пов'язану з завданням, яка не може бути охоплена виключно метою реконструкції, він вводить допоміжне завдання передбачення винагороди для автокодувальника.

Зокрема, окреме оновлення візуальних представлень і динаміки шляхом повторення ітераційних процесів навчання автокодувальника зі згортковим маскуванням ознак і прогнозування винагороди, а також навчання моделі латентної динаміки, яка передбачає візуальні представлення від автокодувальника. Тут підкреслюється, що структура не є схемою попереднього навчання та тонкого налаштування, а постійно оновлює автокодер і модель прихованої динаміки за допомогою зразків, зібраних із взаємодії середовища. Схеми представлені на рисунку 2. [4]

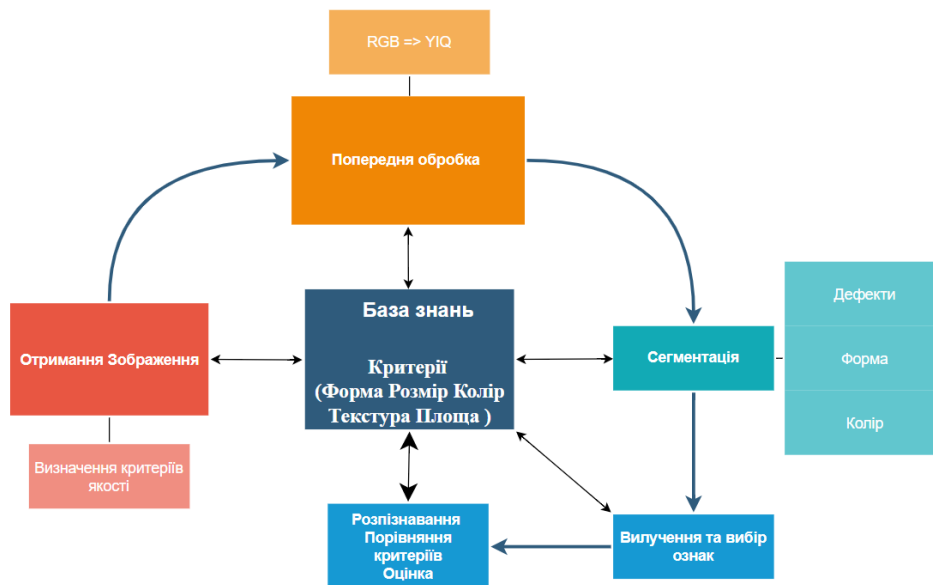


Рисунок 2 – Схема основних етапів комп’ютерного зору для обробки цифрових зображень

4.1. Отримання зображення.

Перш ніж розпочати будь-яку обробку зображень, зображення має бути зроблено камерою [4]. Таким чином, щоб отримати цифрове зображення, необхідні сенсор зображення та здатність оцифровувати сигнал, що виробляється ним. Сенсором може бути телевізійна камера, камера лінійного сканування, відео, сканер і т. д. Якщо вихід сенсора не цифровий, то для оцифрування зображення необхідний аналого-цифровий перетворювач. Цифрове зображення виходить в результаті вибірки і квантування аналогового зображення або створюється вже в цифровому вигляді.

Зазвичай цифрове зображення представляється у вигляді двовимірної матриці дійсних чисел. Конвенція $f(x,y)$ використовується для посилання на зображення розміром $M \times N$, де x позначає номер рядка, а y — номер стовпця. Значення двовимірної функції $f(x, y)$ у будь-якому пікселі з координатами (x_0, y_0) , що позначається $f(x_0, y_0)$, називається інтенсивністю або рівнем сірого зображення в цьому пікселі. З огляду, що зйомка об’єкту ведеться з невеликої відстані при хорошому освітленості та відсутні вагомні шуми, ми не потребуємо високошвидкісної камери з високим розширенням, відповідне зображення можна отримати на ширший спектр камер. Що часто не враховується в інших системах комп’ютерного зору, тим самим перевантажуючи систему.

Попередня обробка. Після отримання цифрового зображення можна застосувати кілька методів попередньої обробки, щоб покращити дані зображення перед обробкою [4]. На цьому етапі зображення обробляється та перетворюється у форму, придатну для подальшого аналізу. Більшість програм комп’ютерного зору вимагають ретельного проектування етапу обробки для досягнення прийнятних результатів. Операції попередньої обробки також називаються фільтрацією.

Тому на цьому етапі пропонується включити згладжування, корекцію експозиції та балансування кольору, шумозаглушення, збільшення різкості, поділ площини зображення, нормалізацію тощо. Що важливо вказати, саме підготовка даних, тобто попередня обробка, являється ключовим фактором не лише у системах комп’ютерного зору, а й у сфері машинного навчання загалом. Також пропонується варіації представлення зображення, наприклад, можна зменшити розширення вдвічі, або ж розвернути по горизонталі або вертикалі. Тобто тут наше завдання отримати чим краще підготовлені дані, тобто зображення, фрукту.

Зображення, отримане після цього етапу, є вхідним для етапу сегментації.

4.2. Сегментація.

Сегментація — це процес поділу цифрового зображення на кілька сегментів (наборів пікселів, також відомих як суперпікселі). Мета полягає в тому, щоб спростити і змінити уявлення зображення на щось, що є більш значущим і легшим для аналізу. Процес грубої сегментації затримує результат задовільного вирішення проблеми DIP. З іншого боку, слабкий процес сегментації в більшості випадків призведе до помилок[5].

В результаті процесу сегментації отримують необроблені піксельні дані, які утворюють межі між регіонами, або інформацію про те, який піксель до якого регіону належить. Сегментація зображення — це операція, яка знаменує перехід від низькорівневої обробки зображення до середнього рівня.

Тут пропонується використання методу порогового визначення та визначення контуру.

4.4. Функція ознак

Після того, як зображення було сегментовано, можна описати отримані окремі області [6]. Вилучення ознак, яке також називають представленням та описом зображення, — це операція, яка виконується для вилучення та виділення ознак з певною кількісною інформацією, яка є важливою для відрізнення одного класу об’єктів від іншого. Це важливий крок у більшості рішень комп’ютерного зору, оскільки знаменує перехід від графічного представлення даних до неграфічного.

Так у роботі для збереження характеристик, витягнутих з об’єкта, створюється масив $n \times 1$, який називається вектором ознак. Вектор ознак — це компактне представлення зображення, його вміст може бути символічним, числовим або обома. Основна проблема на цьому кроці полягає в тому, що виділені ознаки мають бути

незмінними до змін обертання, масштабу, трансляції та контрасту. Отримання інваріантів гарантує, що система комп'ютерного зору зможе розпізнавати об'єкти, навіть якщо вони з'являються з різним контрастом, розміром, положенням і кутом на зображенні.

До важливих вилучених ознак належать точки, прямі лінії, області зі схожими властивостями, колір, текстури, форми та комбінації. Тобто саме те, на основі чого ми можемо судити про якість об'єкту. Для вирішення проблеми інваріантності часто використовуються такі граничні дескриптори, як статистичні моменти або дескриптори Фур'є.

4.4. Розпізнавання.

Цей етап являє собою високий рівень обробки зображень [6]. Розпізнавання — це процес присвоєння мітки об'єкту на основі інформації, наданої його дескрипторами. Більше того, це передбачає присвоєння значення набору розпізнаних об'єктів. Алгоритми розпізнавання аналізують числові властивості різних ознак зображення та класифікують дані за категоріями. Цей етап складається з двох етапів: навчання та тестування.

База знань. Априорні знання про конкретну проблему обробки зображень кодуються у формі бази даних [7]. База даних може бути такою ж простою, як деталізація областей зображення, де відомо, що розташована інформація, що представляє інтерес, що обмежує пошук, який необхідно проводити для пошуку інформації; або може бути досить складним, наприклад, взаємопов'язаний список усіх основних можливих дефектів у проблемі перевірки матеріалу. База знань керує роботою кожного модуля обробки, а також контролює взаємодію між усіма модулями.

5. Визначення критеріїв якості та проведення експерименту

Як було зазначено у вищезгаданому пункті, більшість систем комп'ютерного зору включають попередньо зазначені етапи, але відповідно до специфіки можливі різні варіанти. Так для розробки системи було проведена наступна робота.

Визначення критеріїв якості.

Гібридна система пропонує можливість інтегрувати два або більше представлення знань певної області в одну систему. Однією з головних цілей є отримання додаткових знань, які дозволяють підвищити ефективність глобальної системи. Конкретним прикладом гібридної системи є так звана NSHS, яка в основному заснована на символічному представленні об'єкта, отриманого від людини-експерта, у вигляді правил і системи комп'ютерного зору для отримання числової інформації.

Критерії якості для оцінки груш були отримані при прямій візуальній оцінці експерта з класифікації фруктів на основі його власного досвіду. У цьому випадку категорія присвоюється залежно від значення зовнішніх атрибутів. Існує чотири категорії: категорія екстра, категорія I, категорія II і категорія III. У роботі оцінюється тільки категорія екстра, за якою груша може належати до одного з двох класів: хорошої або поганої якості. На рис.3 зображено приклад груш, які ми будемо кваліфікувати як хорошої та поганої якості, тобто друга груша є з явним візуальним дефектом. Крім того, у таблиці 1 показано підсумок зовнішніх атрибутів груші з пов'язаною назвою змінної, а також значенням і типом.

Отримання зображення.

Для задачі категорії груша було отримано 148 зображень за допомогою цифрової камери. Для повного набору зображень була виконана операція обертання на 90 і 180 за годинниковою стрілкою, подвоєння і зменшення масштабування і додавання шуму. Наприкінці цього процесу набір із 148 груш було розділено на дві категорії: поганої (74) та хорошої (74) якості. На рис.4 зображено приклад різних груш після зміни обертання та масштабу.

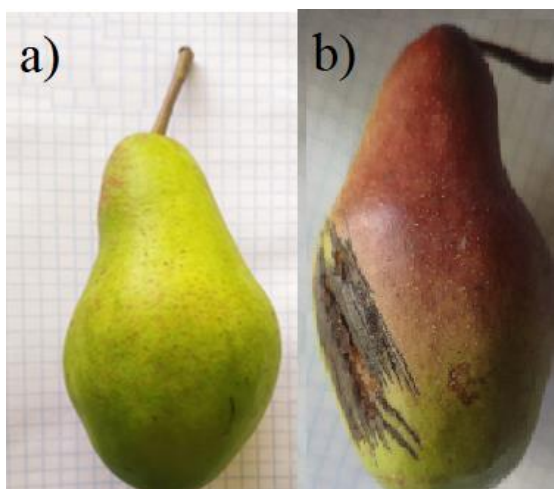


Рисунок 3 - Приклади груші, а) хорошої якості та б) поганої якості з явним візуальним дефектом

Таблиця 1 Критерії, визначені у безрозмірних одиницях для встановлення хорошої та поганої якості, отримані в консультації з відповідними експертами з даної сфери

Атрибут	Акронім	Тип	Значення
Подовжений дефект	LD	Діапазон	0–6
Плямистий дефект	SD	Діапазон	0–2.5
Різні дефекти	VD	Діапазон	0–5
Стебло	S	Двійковий	True/false
Червоний колір	RC	Діапазон	0–255
Зелений колір	GC	Діапазон	0–255
Синій колір	BC	Діапазон	0–255

Попередня обробка зображень. Етап складається з перетворення зображення з колірної моделі RGB в колірну модель YIQ (Яскравість, Інфаза, Квадратура). Основною причиною такого перетворення є полегшення вилучення ознак зображення. Модель YIQ була розрахована, щоб відокремити колір від компонента яскравості через здатність зорової системи людини сприймати більше змін відбиття, ніж зміни відтінку або насиченості. Основна перевага моделі полягає в тому, що коефіцієнт відбиття (Y) та інформація про колір (I та Q) можуть оброблятися окремо. Коефіцієнт відбиття пропорційний кількості світла, що сприймається людським оком.

Вилучення ознак груші. Характеристики кожного зображення були отримані на основі інформації, визначеної людиною-експертом у формі правил та за допомогою обробки зображень у вигляді числових даних. Ці два типи інформації про знання були об'єднані, щоб отримати загальне уявлення про грушу.

Кількість правил, визначених експертами, становила чотири, прикладом одного правила є наступне: «Якщо груша має відповідний колір, має ніжку, має подовжені дефекти, які не перевищують 2 см, і має кілька дефектів, які не перевищують 1 см², та мають плямисті дефекти, які не перевищують 1/4 см², тоді груша відноситься до категорії екстра з хорошою якістю.

Наприкінці цього етапу правила були складені за допомогою штучної нейронної мережі, заснованої на наявних методах (KBANN), щоб отримати представлення інформації, яку в подальшому можна об'єднати з числовими результатами, отриманими з системи комп'ютерного зору. Комбінація була зроблена за допомогою методу під назвою Neusim, який заснований на алгоритмі Фальмана каскадної кореляції.

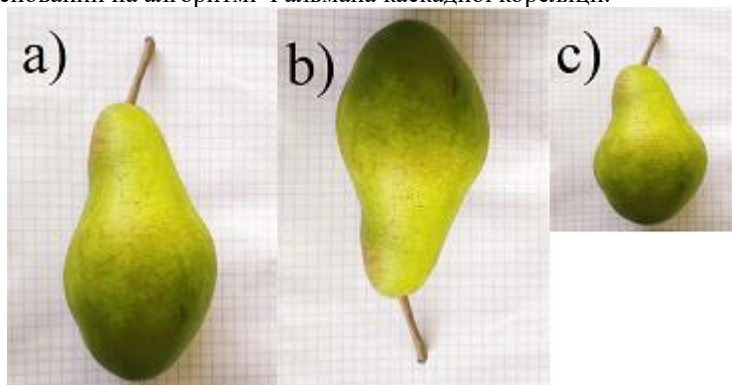


Рисунок 4 - Груша хорошої якості. а) Оригінальне зображення, б) Повернуто на 180 за годинникову стрілку, і в) Зменшено початковий розмір вдвічі

Класифікація груші. Результатом етапу виділення ознаки було об'єднане представлення символічного та числового представлення. Для подальшої класифікації необхідне уточнення цих даних. Це уточнення виконується шляхом повторного запуску методу Neusim, але тепер не для об'єднання знань, а для використання його як класифікатора.

Основною перевагою використання алгоритму Neusim є те, що можна побачити кількість прихованих одиниць, доданих під час процесу навчання, це дуже корисно для моніторингу повного процесу поступового навчання. Результатом цього етапу є рішення про якість груші одного з двох класів, поганого чи хорошого.

Експерименти та результати.

Із загального набору 148 зображень 74 - використано на етапі навчання та 74 - на етапі розпізнавання. Для проведення експериментів було обрано три різні підходи: (а) коннекціоністський підхід, який використовує лише дані, отримані від системи комп'ютерного зору, (б) символічний підхід, який використовує лише дані, отримані із зібраних правил, і (с) NSHS, який є поєднанням коннекціоністського та символічного підходу.

Для тестів із застосуванням коннекціоністського підходу було визначено три сценарії: (а) з використанням числових даних, отриманих із загальних 148 зображень (100 %), (б) з використанням даних, отриманих лише з 111 зображень (75 %), і (с) використання лише даних із 74 зображень (50 %). Для отримання результатів, що стосуються тестового випадку, використовувалися три правила за допомогою символічного підходу. Перше правило під назвою 7 включає наступні сім атрибутів: LD, SD, VD, S, RC, GC і BC. Друге правило під назвою R5 розглядає наступні п'ять атрибутів: RC, GC, BC, S і LD. Нарешті, третє правило під назвою R4 включає наступні чотири атрибути: LD, SD, VD і S. Для випадку підходу NSHS була зроблена комбінація коннекціоністського та символічного. Три правила R7, R5 і R4 були об'єднані зі 100, 75 і 50 % від загальної кількості прикладів. Загальні отримані результати наведені в таблиці 2.

Однією з типових проблем, що спричиняють збої в системах комп'ютерного зору, є відсутність повного опису об'єкта. Це можна спостерігати, перевіряючи результати, за допомогою символічного та коннекціоністського підходів. Цей недолік можна усунути, використовуючи метод для доповнення інформації даними, визначеними знаннями експерта. Системи, які допускають такі типи комбінації, називаються NSHS, як видно з результатів, наведених у таблиці 2; ці системи є дуже ефективними для доповнення необхідних знань для автоматичного огляду об'єктів за допомогою системи комп'ютерного зору.

Наприклад, у чисто символічному підході правила R4 було недостатньо для правильної класифікації, але при його інтеграції з групою числових прикладів (100, 75, 50 %) отримуємо суттєве покращення, оскільки знання, що не містить правила, доповнюється числовою базою прикладів. Практично у всіх випадках досягнуто покращення отримуваних результатів. Подібну систему у перспективі можна залучати для визначення якості більшої номенклатури продукції.

Таблиця 2. Результати досліджень для трьох підходів стосовно випробувань

Підхід	Складені правила	% використаних прикладів	Точність (%)
Коннекціоніст	–	100	95.14
	–	75	91.21
	–	50	90.54
Символічний	R7	–	93
	R5	–	90.12
	R4	–	14.19
NSHS	R7	100	96.62
	R7	75	95.27
	R7	50	90.54
	R5	100	95.27
	R5	75	95.94
	R5	50	96.62
	R4	100	91.22
	R4	75	93.24
	R4	50	94.59

6. Висновки

1. Для створення сталого оцінювального середовища в процесі перевірки якості фруктів запропоновано систему перевірки якості на основі комп'ютерного зору та розроблено навчальну модель відтворення зображення за допомогою штучної нейронної мережі для якості фруктів.

2. У ході перевірки результатів за допомогою символічного та коннекціоністського підходів можна спостерігати збої в системах комп'ютерного зору, зумовлену недостатньо повним описом об'єкта дослідження. Цей недолік можна усунути, за допомогою методу для доповнення інформації даними, визначеними знаннями експерта, завдяки використанню систем, які допускають такі типи комбінації, як NSHS, з підвищенням правильності прийняття рішень до (90 + 6) %.

Посилання

[1] Jorge Luis García-Alcaraz, Aidé Aracely Maldonado-Macías, Guillermo Cortes-Robles. Lean Manufacturing in the Developing World, 117(8), 827–891. 2014, Доступно: <https://link.springer.com/book/10.1007/978-3-319-04951-9>

- [2] Sullivan, W., McDonald, T., & Van Aken, E. (2002). Equipment replacement decisions and lean manufacturing. *Robotics and Computer Integrated Manufacturing*, 18(3–4), 255–265. Доступно: <https://www.sciencedirect.com/science/article/abs/pii/S0736584502000169>
- [3] Choi, J., Ro, Y., & Plataniotis, J. (2011). A comparative study of preprocessing mismatch effects in color image based face recognition. *Pattern Recognition*, 44(2), 412–430. Доступно: <https://www.sciencedirect.com/science/article/abs/pii/S0031320310004061>
- [4] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015. 1, 2. Доступно: <https://ieeexplore.ieee.org/document/7410526>
- [5] Younggyo Seo, Danijar Hafner, Hao Liu, Fangchen Liu, Stephen James, Kimin Lee, Pieter Abbeel. Masked World Models for Visual Control. 2022. 2-3. Доступно: <https://sites.google.com/view/mwm-rl?pli=1>
- [6] David Forsyth, Jean Pons. (2004). *Computer Vision: A Modern Approach*, 928. Доступно: <https://eclass.teicrete.gr/modules/document/file.php/TM152/Books/Computer%20Vision%20-%20A%20Modern%20Approach%20-%20D.%20Forsyth,%20J.%20Ponce.pdf>
- [7] Li, Q., Wang, M., & Gu, W. (2002). Computer vision based system for apple surface defect detection. *Computers and Electronics in Agriculture*, 36(2–3), 215–236. Доступно: <https://www.sciencedirect.com/science/article/abs/pii/S0168169902000935>
- [8] Santos, J., & Rodrigues, F. (2012). Applications of computer vision techniques in the agriculture and food industry: A review. *European Food Research and Technology*, 5(6), 989–1000. Доступно: <https://link.springer.com/article/10.1007/s00217-012-1844-2>