

## MACHINE VISION - BASIC PRINCIPLES, TECHNOLOGIES, APPLICATIONS

BORISLAV T. BORISOV

**ABSTRACT:** Machine vision is an area of artificial intelligence and computer science that deals with the acquisition, analysis, and interpretation of visual information from the environment. This article examines the history, basic principles and technologies, and applications of machine vision in various fields. Key concepts such as image capture, preprocessing, segmentation, feature extraction, and classification are covered. In addition, the paper discusses the use of deep neural networks in machine vision systems, such as CNNs, RNNs, and GANs, and their optimization. Applications of machine vision in medicine and manufacturing, commerce, and agriculture are reviewed, highlighting how these technologies are revolutionizing processes in these fields.

**KEYWORDS:** machine vision, image processing, deep neural networks, segmentation, classification, medical diagnostics, manufacturing automation, neural networks, CNN, RNN, GANs.

**DOI:** <https://doi.org/10.46687/PZKA7056>

## МАШИННО ЗРЕНИЕ - ОСНОВНИ ПРИНЦИПИ, ТЕХНОЛОГИИ, ПРИЛОЖЕНИЯ \*

БОРИСЛАВ Т. БОРИСОВ

**АБСТРАКТ:** Машинното зрение е област от изкуствения интелект и компютърните науки, която се занимава с придобиването, анализа и интерпретацията на визуална информация от околната среда. Тази статия разглежда историята, основните принципи и технологии, както и приложенията на машинното зрение в различни области. Разгледани са ключови концепции като заснемане на изображения, предварителна обработка, сегментация, извличане на характеристики и класификация. В допълнение, статията обсъжда използването на дълбоки невронни мрежи в системите за машинно зрение, като CNN, RNN и GANs, и тяхната оптимизация. Разгледани са приложения на машинното зрение в медицината и производството, търговия, селско стопанство, като се подчертава как тези технологии революционизират процесите в тези области.

### 1 Въведение

Машинното зрение (Computer Vision) е област от изкуствения интелект и компютърните науки, която се занимава с това как компютрите могат да придобиват, анализират и интерпретират визуална информация от околния свят. То съчетава методи и техники за обработка на изображения с алгоритми за извличане на значима информация от тях.

Машинното зрение възниква през 60-те години на XX век с разработването на първите алгоритми за обработка на изображения, като първоначалната цел е била да се

---

\* Настоящата статия е частично финансирана от фонд „Научни изследвания“ на Шуменски Университет „Епископ К. Преславски“ по проект № РД-08-107/30.01.2024 г

автоматизира разпознаването на прости обекти и да се имитират основните човешки зрителни функции. Първите опити, от които започват да се разработват базови техники за сегментация и идентификация на обекти, са извършени в лаборатории на големи университети. Експериментите започват през 1959 г., когато неврофизиолозите Дейвид Хюбел (David Hubel) и Торстен Визел (Torsten Wiesel) показват набор от изображения на котка, в мозъка на която са поставени електроди, регистриращи невронната активност. Анализирайки реакциите в мозъка ѝ, откриват, че той реагира първо на твърди ръбове или линии достигайки до до изводи, че обработката на изображения започва с прости форми като прави ръбове [1].

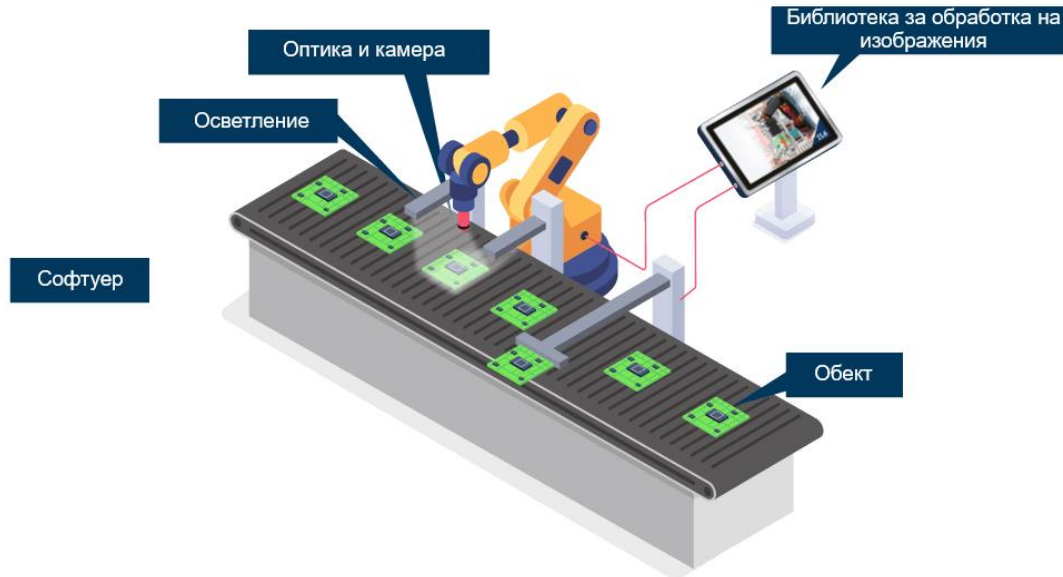
Алгоритмите за обработка на изображения започват да се подобряват през 70-те години на миналия век и се появяват по-сложни методи за разпознаване на граници и контури. През този период машинното зрение започва да намира първите си приложения в роботиката и индустриалната автоматизация. За един от основателите на машинното зрение е считан Лари Робъртс (Lawrence G. Roberts) с публикуваната през 1963г докторска дисертация „Машинното възприемане на триизмерни тела“ [2], в която описва процеса на извличане на 3D информация за твърди обекти от 2D снимки, като основно свежда визуалния свят до прости геометрични форми. Дейвид Мар (David Marr) е пионер в теоретичните основи на зрението, Мар формулира тристепенна теория за компютърното зрение: суровите данни, 2.5D скица и пълно възприемане на сцена. Той изследва как мозъкът интерпретира визуалните данни и разработва модели, които вдъхновяват много бъдещи изследвания в областта [3].

След 1980 развитието на компютърната графика и първите по-сериозни процесори правят възможно по-бързото обработване на изображения. Технологиите за заснемане, като цифровите камери, също се развиват, което отваря нови хоризонти пред машинното зрение. Един от водещите изследователи в областта на компютърното зрение и роботиката е Такео Канаде (Takeo Kanade), който разработва множество методи за проследяване на движения, разпознаване на лица и възстановяване на 3D обекти [4]. Неговата работа се използва широко в индустрията и научните среди.

В началото на XXI век започва бурно развитие на дълбокото обучение (deep learning) и невронните мрежи, машинното зрение навлиза в нова ера. Въвеждането на големи набори от данни, като ImageNet, и наличието на мощни графични процесори (GPUs) позволяват по-сложни и ефективни модели, които могат да разпознават сложни обекти и сцени с много висока точност. Ян Лекун (Yann LeCun) е изследовател, известен с разработката на конволюционни невронни мрежи (CNNs), които играят ключова роля в модерното машинно зрение. Неговите изследвания в областта на разпознаването на изображенията и дълбокото обучение са част от основата на съвременните приложения на машинното зрение [5].

## 2 Машинно зрение: основни принципи и технологии

### 2.1 Принципна схема [6]



Фигура 1 – Принципна схема на система с машинно зрение

На *Фигура 1* е показана принципната схема на машинното зрение обикновено включва следните компоненти :

#### 2.1.1 Източник на осветление

Правилното осветление е критично за качеството на изображението и обработката му. Осветлението може да бъде дифузно или насочено, в зависимост от задачата. Например, за инспекция на детайли може да се използва определен вид осветление, което подчертава дефекти.

#### 2.1.2 Камера и оптика

Камерата и нейната оптика са отговорни за улавяне на изображението. Изборът на подходящи обективи, резолюция и скорост на заснемане са ключови за постигането на качествени данни. Често се използват индустриални камери с висока резолюция.

#### 2.1.3 Сензор за изображения

Сензорът преобразува светлината, която камерата улавя, в цифров сигнал. Съществуват различни видове сензори, като **CCD** (Charge-Coupled Device) и **CMOS** (Complementary Metal-Oxide-Semiconductor), които се използват в зависимост от приложението.

#### 2.1.4 Обработка на изображения

След заснемането на изображението, системата за машинно зрение извършва предварителна обработка, за да го подготви за анализ. Това може да включва техники като филтриране на шума, регулиране на контраста и изостряне на контурите.

#### 2.1.5 Анализ и алгоритми

В основата на машинното зрение са алгоритмите, които извършват анализа на изображенията. Може да се използват различни методи за разпознаване на форми, текстури и обекти, като по-напредналите системи разчитат на машинно обучение и дълбоки невронни мрежи за вземане на решения на база визуални данни.

#### 2.1.6 Изход и интерфейс

След като анализът е извършен и е взето решение, системата предава резултатите към друг софтуер или хардуер. В индустриални приложения това може да бъде сигнал към робот или производствена линия, които предприемат действия въз основа на визуалния анализ.

## 2.2 Основни технологии

Машинното зрение се основава на няколко ключови принципа, които му позволяват да възприема, обработва и анализира визуална информация. Основната цел е компютърните системи да "разбират" образи по начин, подобен на човешкото зрение, но чрез алгоритми и обработка на данни. Принципната схема на машинното зрение обикновено включва няколко етапа, всеки от които допринася за успешното изпълнение на конкретни задачи, като разпознаване на обекти, класификация, детекция на дефекти и т.н.

Основните принципи на машинното зрение са

#### 2.2.1 Заснемане на изображения

Първият етап в машинното зрение е придобиването на визуална информация, което обикновено се извършва чрез цифрови камери, сензори или други устройства, които преобразуват физическите визуални сигнали в цифров формат. Това включва както 2D (двуизмерни) изображения, така и 3D (триизмерни) данни от технологии като LIDAR или стереоскопично зрение.

#### 2.2.2 Предварителна обработка на изображението

След заснемането на изображенията, те трябва да бъдат подготвени за анализ. Това включва различни техники за обработка на изображения, които помагат за подобряване на качеството на данните. **Филтрирането** премахва шум или нежелани артефакти от изображението. Посредством **контрастно усилване** се подобряват разликите между обектите и фона, за да се улесни по-късното разпознаване. с **нормализиране** се преобразуват изображенията към определен формат, мащаб или цвят схема.

#### 2.2.3 Сегментация

След обработката, изображението се сегментира, което означава, че се разделя на области или обекти, които могат да бъдат анализирани отделно. Сегментацията е важна, тъй като изолира отделни обекти или области от изображението, които представляват интерес. Обикновено се използват техники като **откриване на ръбове (edge detection)** за намиране на граници между различни обекти или части от изображението, както и **сегментиране по цвят или текстури** за разделяне на изображението въз основа на еднаквост в цветовете или текстурата.

#### 2.2.4 Извличане на характеристики

Този етап включва идентифициране на важни характеристики или "черти" (features) от изображението. Характеристиките могат да включват форми и

контури на обекти, цветове и текстури, както и ключови точки като ъгли и ръбове (например алгоритми като SIFT или SURF). Извлечените характеристики представляват информация, която алгоритмите за машинно зрение използват за разпознаване на обекти или за класификация.

#### 2.2.5 Класификация и разпознаване на обекти

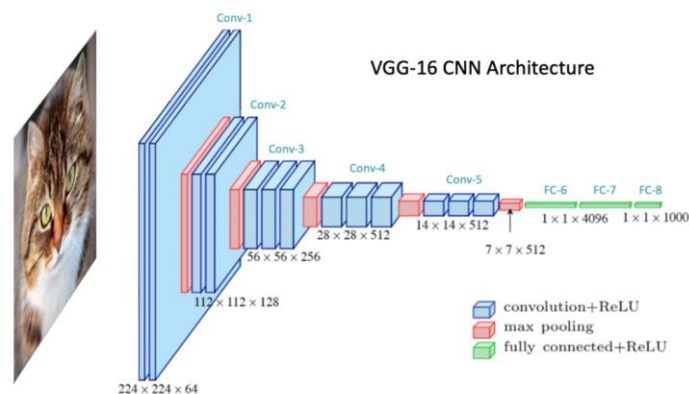
След извлечането на характеристиките се пристъпва към разпознаването на обекти или класификацията им. В този етап се прилагат различни алгоритми за машинно обучение (особено дълбоки невронни мрежи или конволюционни невронни мрежи), които могат да разпознават и класифицират обекти въз основа на обучени модели.

#### 2.2.6 Вземане на решение

Последният етап в схемата на машинното зрение е вземането на решение. На база на резултатите от анализа и класификацията, системата взема решение за последващи действия. Например, в производствената линия, системата може да реши дали даден продукт е дефектен или не, или в автономно превозно средство може да идентифицира препятствие и да избегне сблъсък.

### 2.3 Невронни мрежи в машинното зрение и тяхната оптимизация

Машинното зрение е пряко свързано с използването на невронни мрежи, като особено внимание се обръща на дълбоките невронни мрежи (Deep Neural Networks - DNN), които имат голямо значение за постигането на висока точност в задачите за разпознаване на обекти, класификация и сегментация. Невронните мрежи играят решаваща роля в машинното зрение и предлагат възможности за автоматизация и оптимизация на процеси. Чрез различни техники, като пренос на обучение, квантизация и регуларизация, тези модели могат да бъдат адаптирани за индустриални приложения с високи изисквания за производителност и ефективност. Основните видове невронни мрежи, използвани в машинното зрение:



Фигура 2 – Мрежова архитектура на CNN

#### 2.3.1 Конволюционни невронни мрежи (Convolutional Neural Networks - CNNs)

Това са основният модел (Фигура 2 [15]), използван в машинното зрение и са особено ефективни при задачи за разпознаване на изображения, защото използват специализирани слоеве, наречени конволюционни слоеве (convolutional layers), които извличат характеристиките на изображенията, като

ръбове, текстури и обекти. Основни характеристики на CNN са **конволюционни филтри**, които сканират изображенията, за да открият специфични елементи и структури, **pooling слоеве** за намаляване размерите на данните и помагат да се избегне пренасяне с информация, като същевременно запазват важните характеристики и **пълносвързани слоеве** за вземане на окончателни решения, като класификация на изображенията.

### 2.3.2 Рекурентни невронни мрежи (Recurrent Neural Networks - RNNs)

Рекурентните невронни мрежи (RNN) са по-често използвани в задачи, свързани с последователни данни, като текст или аудио, но те намират приложение и в машинното зрение. Използват в задачи, където се изисква обработка на последователности от изображения (напр. видеоклипове). Дългосрочна и краткосрочна памет (LSTM) и GRU (Gated Recurrent Units) са подобрени версии на RNN, които могат да се използват в анализ на времеви серии на изображения.

### 2.3.3 Генеративни състезателни мрежи (Generative Adversarial Networks - GANs)

Генеративните състезателни мрежи (GAN) представляват състезание между два модела – генератор и дискриминатор. Те са мощен инструмент за създаване на реалистични изображения и симулиране на визуална информация. В производството, например, GANs могат да се използват за симулиране на дефекти за тренировка на модели, които да разпознават дефектни продукти.

### 2.3.4 Мрежи за сегментация (Fully Convolutional Networks - FCNs)

FCN мрежите са специализирани за задачи за сегментация, където всяка точка в изображението трябва да бъде класифицирана. Тези мрежи са широко използвани за сегментация на обекти или анализ на дефекти в производствени среди.

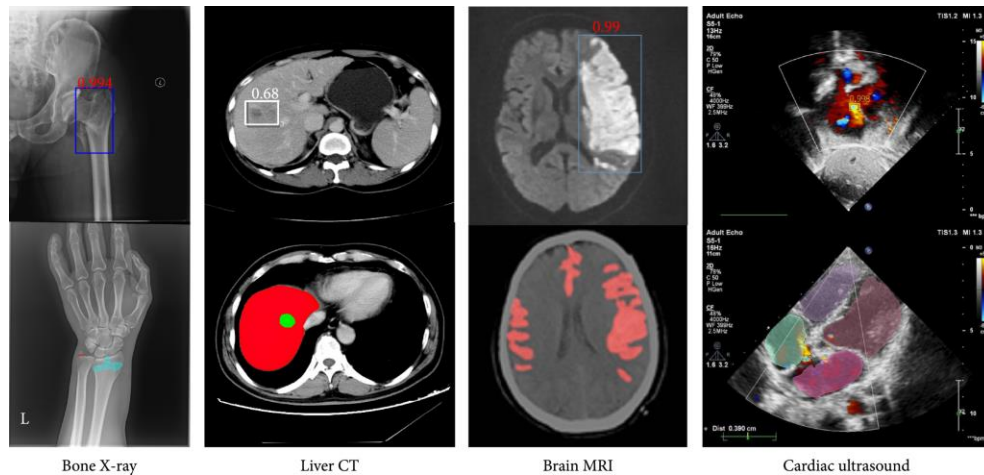
Невроните мрежи, използвани в системите с машинно зрение, изискват голям изчислителен ресурс при тяхното адаптиране и обучение. Съществуват различни начини за оптимизация на невронните мрежи, като целта е да се увеличи бързодействието и да се намали обемът от данни, необходими за обучение. Оптимизацията с пренос на обучение (Transfer Learning) е особено ефективен подход, защото предварително обучена мрежа, като ResNet [7] или VGG [8], се използва за решаване на нова задача с малко количество нови данни. Предварително извлечените характеристики спестяват време за обучение и намаляват нуждата от голям набор от данни. Квантизацията е процес на намаляване на прецизността на числата, които се използват в модела (напр. замяна на 32-битови с 8-битови числа), което води до намаляване на размера на модела и по-бързо изчисление с минимална загуба на точност. Това е полезно в индустриални приложения, където скоростта и ресурсната ефективност са критични [9]. Използването на специализиран хардуер като графични процесори (GPUs) или тензорни процесори (TPUs) за ускоряване на тренировъчния процес е от съществено значение при работа с големи обеми визуални данни. Тези устройства са оптимизирани за паралелно изчисление, което значително ускорява обработката на изображения [10]. Подходящи за вградени системи и приложения в реално време са по-леки архитектури, като MobileNet и EfficientNet. Тези мрежи предлагат балансирано съотношение между точност и изчислителни ресурси [11].

За да се избегне преобучение и подобряване на обобщаващите способности на модела, могат да се използват различни методи за регуларизация, като отпадане на случаен избор на

неврони (Dropout), които се игнорират по време на обучението [12]. L2-регуляризацията е метод с добавяне на малки ограничения върху стойностите на параметрите на модела, за да се избегне прекомерното му приспособяване към тренировъчните данни [13]. Аугментацията на данни използва на техники за изкуствено разширяване на набора от данни чрез трансформации на изображенията (завъртане, скалиране и др.).

## Машинно зрение – приложения

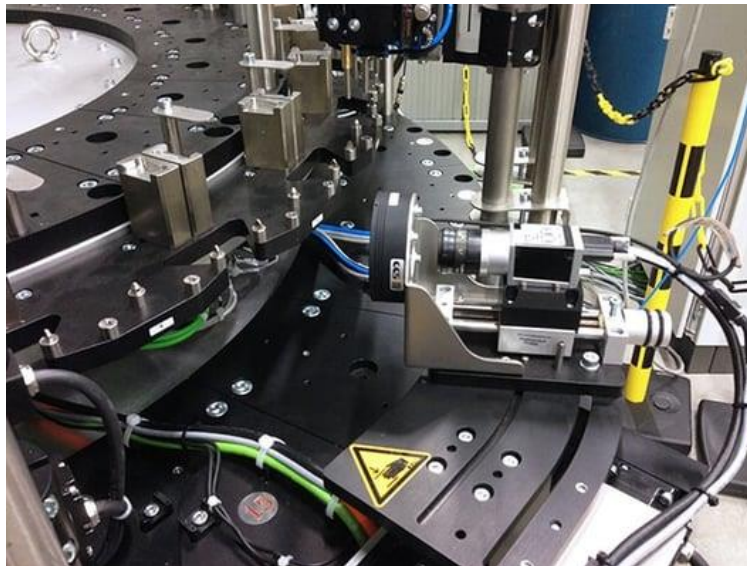
### 2.4 Приложение на машинното зрение в медицината.



Фигура 3 – Приложение на машинно зрение в медицината

Машинното зрение се използва за автоматичен анализ на медицински изображения, като рентгенови снимки, компютърна томография (КТ), магнитен резонанс (MRI) и ултразвукови изследвания (Фигура 3) [16]. Алгоритмите за дълбоко обучение помагат за откриване на аномалии като тумори, лезии или атеросклероза, което подпомага лекарите в диагностицирането на заболяванията. Системи за машинно зрение се прилагат за анализ на мамографии и биопсии, за да откриват ракови клетки в ранен стадий. Това подобрява шансовете за успешно лечение чрез предоставяне на ранни предупредителни сигнали, които може да са незабележими за човешкото око. Роботи, оборудвани с машинно зрение, се използват в хирургията за извършване на минимално инвазивни операции. Те могат да извършват изключително прецизни движения въз основа на визуални данни, предоставяни в реално време, което подобрява безопасността и точността на интервенциите. Машинното зрение се използва за анализ на кожни лезии и образувания, като родилни петна или потенциално ракови изменения. Алгоритмите могат да класифицират измененията по тип и да предоставят информация за тяхната вероятна злокачественост. В болничните заведения системите за машинно зрение могат да проследяват пациенти, за да оценят състоянието им, като наблюдават двигателните им функции или сигнализируют за възможни рискове, като падане от легло. Това е от голямо значение за автоматизацията на някои аспекти на грижите за пациентите.

## 2.5 Приложение на машинното зрение в производството



**Фигура 4 – Приложение на машинно зрение в производството**

Машинното зрение е ключов инструмент за автоматизация на контролно-измервателните процеси и контрол на качеството. Автоматизирано откриване на дефекти чрез анализ на изображения на продукти, машинните системи могат да откриват несъвършенства, които са трудно забележими за човешкото око. Повишаване на точността на инспекцията благодарение на прецизни алгоритми за анализ, системите могат да извършват много по-точна проверка на продукти в сравнение с ръчната инспекция.

Оптимизация на производствените линии (Фигура 4 [16]) използва машинно зрение за мониторинг на процесите в реално време, наблюдение на производствените линии за отклонения в характеристиките на произвежданите артикули, автоматизация на операции като сортиране на компоненти или сглобяване на детайли. Роботите, оборудвани с машинно зрение, могат да извършват сложни задачи: като манипулиране на малки компоненти или сглобяване на продукти с висока точност, да се адаптират към променяща се среда: благодарение на системите за разпознаване и анализ на обекти.

## 2.6 Приложение на машинното зрение в търговията.

Машинното зрение намира широко приложение в автоматизацията на складови операции, като автоматизирано сортиране на стоки: чрез системи за разпознаване на баркодове или етикети, инвентаризация в реално време: чрез използването на камери и сензори за проследяване на наличността на стоки и оптимизиране на складовите запаси. В търговските обекти машинното зрение допринася за ускоряване на процесите на касовите зони чрез автоматизирано разпознаване на продукти: системи за разпознаване на обекти, които ускоряват процеса на сканиране на продукти и намаляват времето за обслужване на клиентите. При обработка на плащания чрез разпознаване на лицето на клиента за по-бързи и сигурни трансакции. В маркетинговия анализ и поведение на потребителите, системите за машинно зрение могат да анализират поведението на клиентите, като проследяват движението им в магазина, времето, което прекарват в различни секции, и предпочитаните продукти. Посредством разпознаване на клиентите и използване на данни за предишни покупки такива системи се използват за предлагане на персонализирани оферти.



## 2.7 Приложение на машинното зрение в селското стопанство.

Машинното зрение може да се използва за наблюдение на здравето на растенията и културите. Дронове и наземни системи, оборудвани с камери и сензори, улавят изображения на полетата, които след това се анализират за откриване на признаци на болести, недостиг на хранителни вещества или вредители. Това позволява на фермерите да реагират своевременно и целенасочено да прилагат торове или пестициди. Машинното зрение се използва за автоматизиране на процесите на сортиране на продукцията по качество, размер и зрялост. Системите могат да разпознават дефекти, като петна или увреждания по плодовете, което улеснява премахването на некачествена продукция.

Камерите и алгоритмите за машинно зрение могат да оценяват броя на плодовете на дървета или растения, като предоставят точни прогнози за очаквания добив. Това улеснява планирането на реколтата и оптимизиране на логистиката. Роботи, оборудвани с машинно зрение, могат да разпознават плевели и автоматично да ги премахват, без да увреждат основните култури. Системи за прибиране на реколтата също използват машинно зрение за разпознаване на зрели плодове и зеленчуци, за да ги събират автоматизирано.

## 4 Заключение

Машинното зрение представлява ключова технология за оптимизацията на процесите както в производството, така и в търговията. Чрез автоматизацията на рутинни операции, контрол на качеството и по-добро управление на ресурсите, тази технология допринася за по-голяма ефективност и конкурентоспособност на компаниите. Въпреки първоначалните разходи и предизвикателствата при внедряване, ползите от машинното зрение го превръщат в основен инструмент за бъдещето на мадицината, индустриалните и търговските сектори.

### ЛИТЕРАТУРА:

- [1] Hubel, D., Wiesel, T. (2005). Brain and visual perception: the story of a 25-year collaboration. Oxford p.106
- [2] Roberts, L. (1963). "Machine Perception of Three-Dimensional Solids" Thesis (Ph. D.)-- Massachusetts Institute of Technology, Dept. of Electrical Engineering
- [3] Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Henry Holt and Co., Inc., USA.
- [4] Kanade, T. (1973) Picture Proceasing System by Computer Complex and Recognition of Human Faces. PhD thesis, Department of Information Science, Kyoto University, November 1973
- [5] Lecun, Y., Bengio, Y., Hinton. G. (2015). Deep learning. Nature, 521 (7553), pp.436-444. ff10.1038/nature14539ff. ffhal-04206682f
- [6] URL <https://www.technexion.com/resources/embedded-vision-vs-machine-vision-everything-you-need-to-know/> TechNexion Ltd. official page (посетена на 10.09.2024)
- [7] Liang, J. (2020). Image classification based on RESNET. Journal of Physics: Conference Series. 1634. 012110. 10.1088/1742-6596/1634/1/012110.

- [8] Zhou, Y., Chang, H., Lu, Y., Lu, X., Zhou, R. (2020). Improving the Performance of VGG Through Different Granularity Feature Combinations. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.3031908.
- [9] Хайров, Э.М., Мальсагов, М.Ю., Карандашев, Я.М., Пушкарёва, М.М., Крыжановский, Б.В. (2020). КВАНТИЗАЦИЯ ВЕСОВ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ПОСЛЕ ОБУЧЕНИЯ. NANOINDUSTRY Russia. 96. 543-548. 10.22184/1993-8578.2020.13.3s.549.555.
- [10] Liu, S., Ju, T. (2024). APapo: An asynchronous parallel optimization method for DNN models, Future Generation Computer Systems, Volume 152, Pages 317-330, ISSN 0167-739X
- [11] Arjun, P.A., Suryanarayan, S., Viswamanav, R.S., Abhishek, S., Anjali, T. (2024). Unveiling Underwater Structures: MobileNet vs. EfficientNet in Sonar Image Detection, Procedia Computer Science, Volume 233, Pages 518-527, ISSN 1877-0509
- [12] Srivastava, N. (2013). Improving Neural Networks with Dropout. Master's thesis, University of Toronto.
- [13] Cortes, C., Mohri, M., Rostamizadeh, A. (2012). L2 Regularization for Learning Kernels. Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, UAI 2009.
- [14] Shorten, C., Khoshgoftaar, T.M. (2019) A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60.
- [15] URL: <https://learnopencv.com/understanding-convolutional-neural-networks-cnn/> LearnOpenCV official page (посетена на 14.09.2024)
- [16] Liu, X., Gao, K., Liu, B., Pan, C., Liang, K., Yan, L., Ma, J., He, F., Zhang, S., Pan, S. (2021) . Advances in Deep Learning-Based Medical Image Analysis. Health Data Sci.
- [17] URL: <https://www.averna.com/en/expertise/vision-optical-inspection> –Averna Ltd. official page (посетена на 15.09.2024)

**Борислав Тодоров Борисов**

Шуменски университет „Епископ Константин Преславски“

[b.t.borisov@shu.bg](mailto:b.t.borisov@shu.bg)