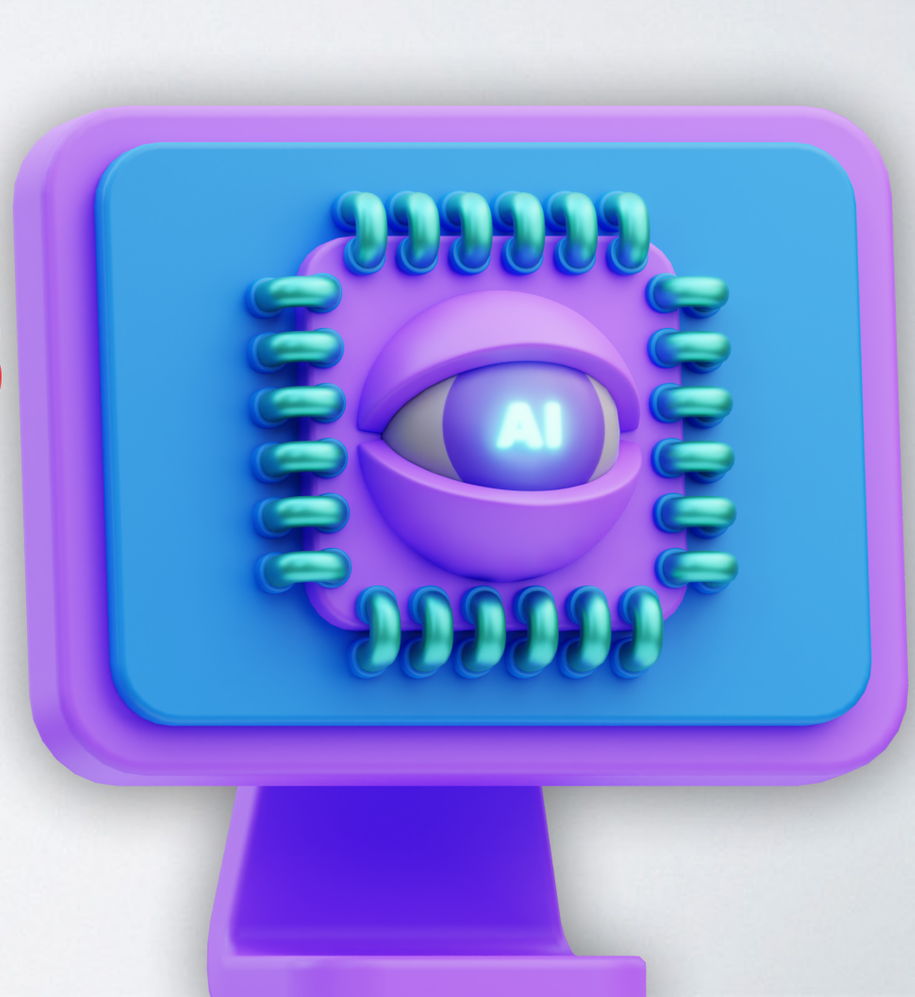
INTRO TO COMPUTER VISION
HISTORICAL TIMELINE OF COMPUTER VISION

2000 - 2020’s



A PRESENTATION BY:

Jeffery Dirden
Bradley Johnson

ITAI 1378
PROFESSOR MCMANUS

Computer vision is used in a variety of different industries including: Retail, E-commerce, Transportation, Manufacturing, Security and Safety, Healthcare, Construction, Gaming, Sports, and Education just too name a few. As time goes on it's persistently evolving and improving. It's helping these industries become more efficient, safer, and smarter. Being able to automate tasks that were previously done manually, it's improving accuracy, and reducing costs to most of these industries. Let's give a couple examples how its impacting industries.

IMPACTFUL APPLICATIONS

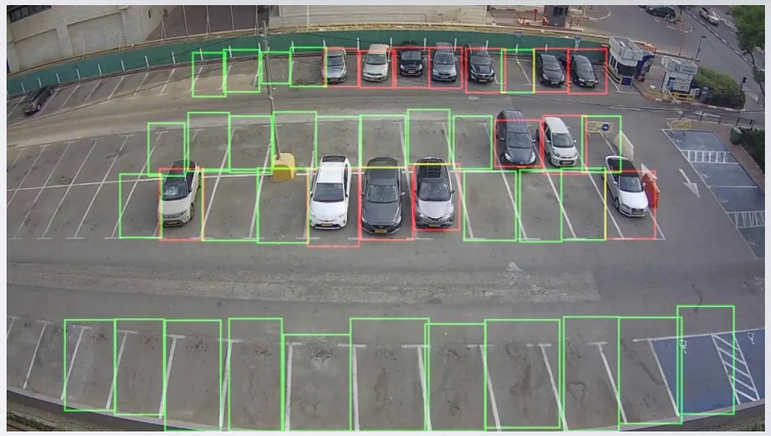
In the Transportation Industry

- **Autonomous Vehicles (Self Driving Cars)**

Autonomous Vehicles are able to detect objects, interpret road signs and markings. They can even make decisions on steering, accelerating and braking using computer vision.

- **Traffic Management**

Applications are used to monitor and manage traffic, including detecting and analyzing congestion, monitoring and managing parking spaces, and identifying and enforcing traffic violations.

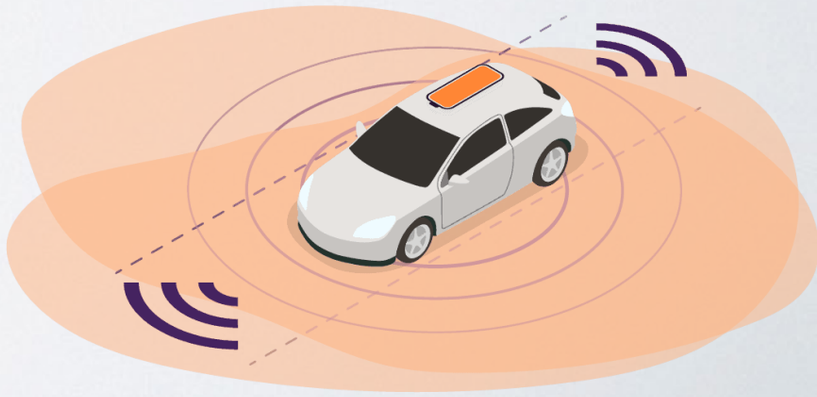


- **Safety Systems**

Safety system applications are used to alert drivers of potential hazards they may encounter on the road. Being able to identify pedestrians, cyclists, other vehicles, and other potential hazards on its own.

- **Airport Security**

Airports use computer vision to identify potential security threats and prohibited items, such as weapons, liquids, explosives, contraband and other threats that may compromise the security and integrity of the establishment.

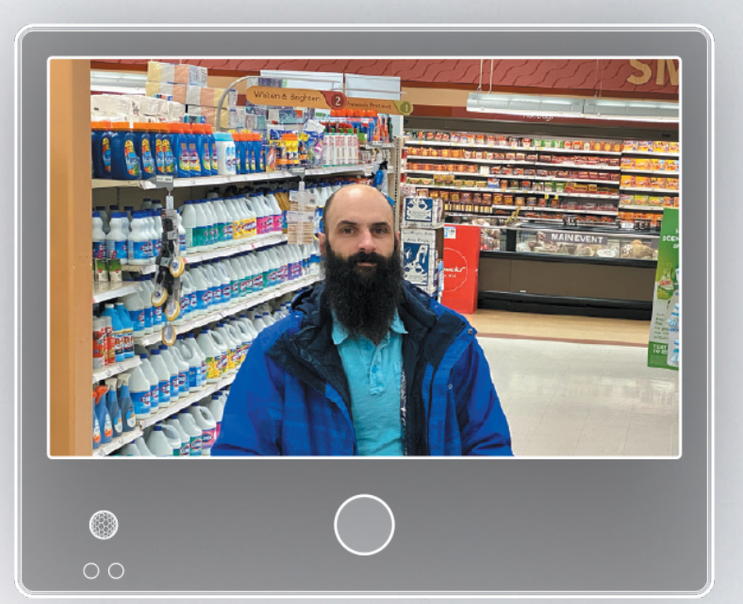


IMPACTFUL APPLICATIONS

In the Retail and E-Commerce Industry

- ### Interaction Free Shopping (Self Checkout)

Self check first appeared in the late 1980's but it would not grow to prominence until the early 2000's. Computer vision is used by allowing cameras and image processing algorithms to automatically detect and track items placed in a customer's shopping cart.



- ### Loss Prevention

Retail stores are using AI vision solutions to monitor shopper activity in a non-intrusive and customer friendly manner. This technology is constantly evolving helping companies protect product by identifying potential thefts in progress.

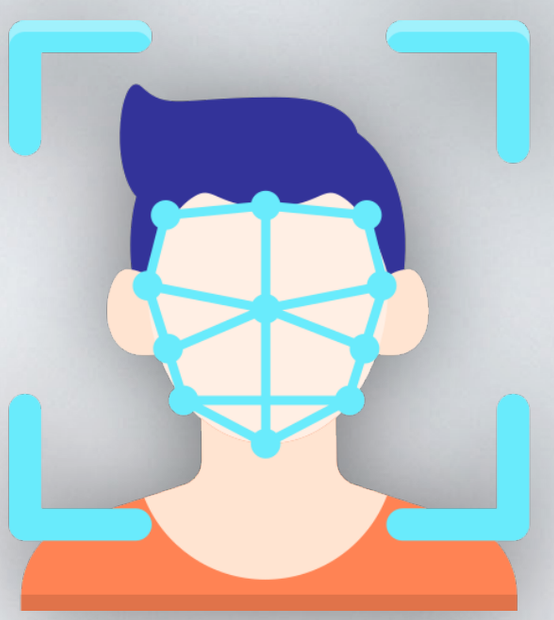
- ### Virtual Mirrors and Recommendation Engines

Virtual mirrors may become the central focus of personalization and customer experience in retail. It's basically a traditional mirror with a display behind the glass. Powered with computer vision cameras and Augmented Reality, these mirrors can display a broad range of information, which, in turn, helps buyers connect with the brand better.

- ### Retail Heat Maps

Retail heat maps can help understand the functionality of the stores and identify customer behavior in them. But what is a heat map? In simple terms, a heat map is just a technical way of saying a shaded matrix where single values in a matrix are displayed as colors.

TIMELINE



2000 - 2004

2001: Eigenfaces for Face Recognition

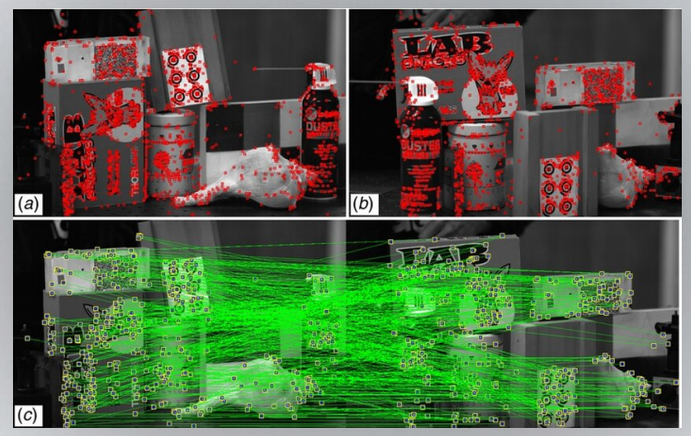
Development of Eigenfaces, a method using principal component analysis (PCA) for face recognition, made important groundwork for the facial recognition technology.

2001: Viola-Jones Object Detection Framework

Paul Viola and Michael Jones introduced a real time object detection framework, known for face detection, which became used due to its efficiency.

TIMELINE

2005 - 2009



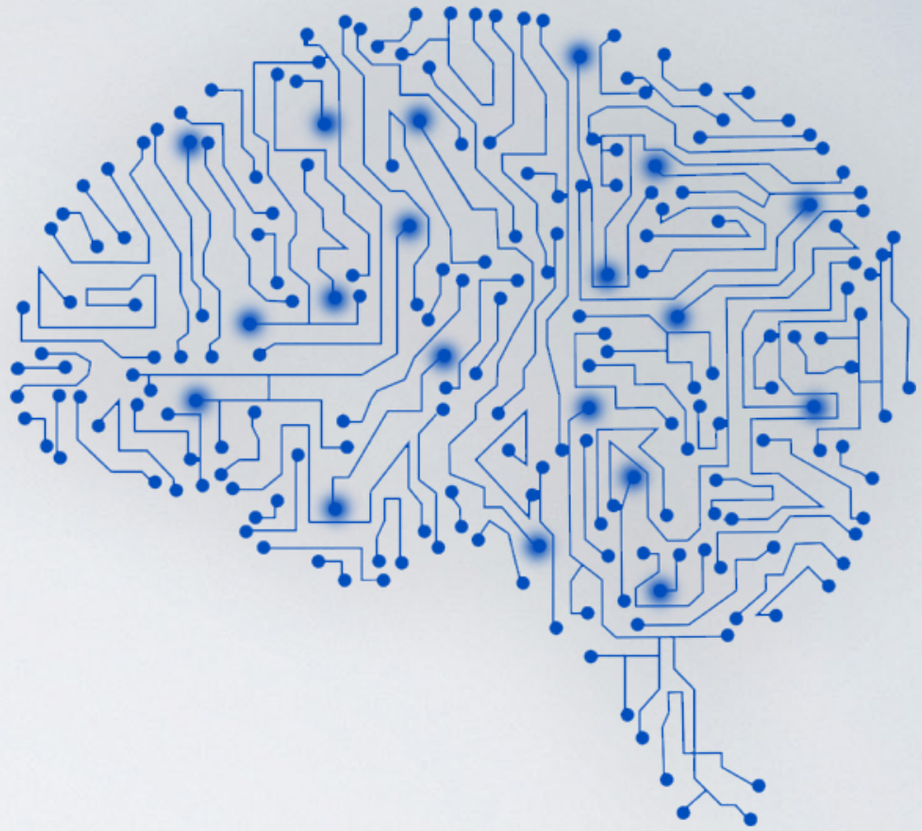
2004-2005: Scale-Invariant Feature Transform (SIFT)

David Lowe's SIFT algorithm provided a method for detecting and describing local features in images, becoming a crucial component in recognition and matching.

2006: Histogram of Oriented Gradients (HOG)

The HOG descriptor, introduced by Navneet Dalal and Bill Triggs, was a popular feature extraction technique. Mainly in human detection tasks.

TIMELINE



2010- 2014

2012: AlexNet and ImageNet

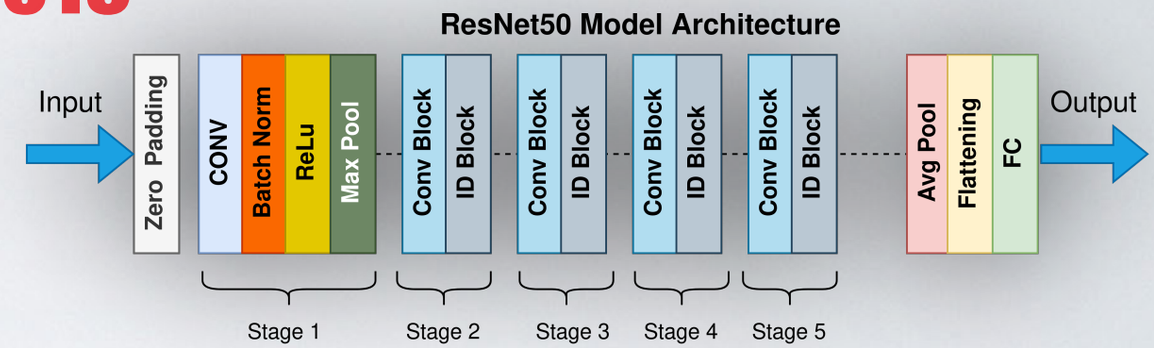
Geoffrey Hinton, Alex Krizhevsky, and Ilya Sutskever introduced AlexNet, a deep convolutional neural network (CNN) that won the ImageNet Large Scale Visual Recognition Challenge in 2012. This solidified the power of deep learning for computer vision tasks, evolutionzing the field.

2013: R-CNN (Region-based Convolutional Neural Networks)

Ross Girshick introduced R-CNN, it significantly improved object detection performance by combining region proposals with CNNs, leading to more accurate detections, classifications of objects within a image.

TIMELINE

2015- 2019



2015: ResNet (Residual Networks)

Kaiming He, and his team developed ResNet, a deep neural network architecture. By solving problems of vanishing gradients in very deep networks, enabling the training of much deeper CNNs. ResNet became foundational in tasks like image classification and object detection.

2015: Faster R-CNN

Ross Girshick and his squad introduced Faster R-CNN, an improvement over R-CNN that included a Region Proposal Network (RPN), significantly speeding up the object detection process.

2016: YOLO (You Only Look Once)

Joseph Redmon introduced YOLO, an object detection model known for its speed and efficiency. Unlike previous methods, YOLO framed object detection as a single regression problem, allowing real-time object detection.

TIMELINE

2020 - Present

DALL·E

2020: Vision Transformers (ViT)

The introduction of Vision Transformers by Google Research marked a shift from CNNs to transformer models for vision tasks. ViTs apply transformer architecture, originally developed for NLP, to image recognition tasks, achieving great results.

2020: CLIP (Contrastive Language–Image Pretraining)

Developed by OpenAI, CLIP demonstrates the ability to perform zero shot learning on various image classification tasks, by training on a wide variety of image text pairs, greatly expanding the versatility of computer vision models.

2021: DALL-E and Image Synthesis

OpenAI introduced DALL-E, capable of generating images from text descriptions. This advancement in AI demonstrated significant progress in the ability to understand and generate visual content based on natural language prompts.

2023: Stable Diffusion

The release of Stable Diffusion, a text to image generation model, brought high quality, generative AI to the mainstream, the integration of computer vision with creative and design industries.

INFLUENTIAL FIGURES

- Takeo Kanade: A professor at Carnegie Mellon University (CMU), was a pioneer in the field of computer vision, focusing in developing real time vision systems. He helped lay the groundwork for many computer vision applications, including autonomous robotics and facial recognition.

- Andrew Blake: A leading researcher in computer vision, Andrew's work on probabilistic models and algorithms for object tracking and image segmentation has had significant impacts on the field. His contributions too active contours and visual motion estimation are essential.

- Pietro Perona: Known for his work in image segmentation and visual recognition, he made substantial contributions to understanding how machines can interpret visual data. His research on anisotropic diffusion and the development of the Viola-Jones object detection framework has been super influential.

INFLUENTIAL FIGURES

Jitendra Malik: A prominent figure in computer vision, Jitendra's research has covered multiple topics, including object recognition, image segmentation, and computational models of human vision. His work on normalized cuts for image segmentation is very notable.

Fei-Fei Li: Co-created the ImageNet dataset, a large scale dataset that has become a cornerstone for training and evaluating computer vision algorithms. Her work has been pivotal in advancing deep learning, & computer vision.

David Lowe: He developed the Scale Invariant Feature Transform (SIFT), a robust algorithm for detecting and describing local features in images. SIFT has been one of the most widely used methods in computer vision for object recognition.

INFLUENTIAL FIGURES

- **Ian Goodfellow: Ian introduced Generative Adversarial Networks (GANs) in 2014, a groundbreaking approach in generative modeling that has had a great impact on computer vision, specifically in the generation and transformation of images.**

- **Ross Girshick: Developed Region based Convolutional Neural Networks (R-CNN), which advanced the field of object detection. His work on Faster R-CNN and Mask R-CNN has become a standard in object detection and instance segmentation.**

- **Pieter Abbeel: Pieter has made big contributions to the integration of deep learning with robotics, specifically in reinforcement learning and the application of vision to robotic manipulation and autonomous systems.**

INFLUENTIAL FIGURES

Ali Farhadi: Has contributed significantly to the fields of object recognition and visual question answering. His work on real time, energy efficient computer vision models, especially through Xnor.ai, has been influential in making AI more accessible for edge devices.

Raquel Urtasun: Is a leading researcher in autonomous driving, known for her work on perception systems that enable self driving cars to interpret and navigate complex environments. Her work combines computer vision, machine learning, and robotics.

Aude Oliva: Her research on understanding visual cognition and developing models that bridge human and machine perception has been instrumental in advancing AI systems that can interpret and generate images in a human like creativity/manner.

Alexei A. Efros: He contributed to several areas of computer vision, including image synthesis, photo manipulation, and scene understanding. His work on neural networks for image translation and generation, has helped push the boundaries of what AI can achieve in visual creativity.

-

KEY ADVANCEMENTS

2000-2004

The Foundations

- **2001: Eigenfaces for Face Recognition:**
  -The development of Eigenfaces, a method using principal component analysis for face recognition, laid important groundwork for facial recognition technology.

- **2001: Viola-Jones Object Detection Framework:**
  -Paul Viola and Michael Jones introduced a real time object detection framework, especially known for face detection, which became widely used due because of its efficiency.

KEY ADVANCEMENTS

2005-2009

Feature Detection and Early Machine Learning

- **2004-2005: Scale-Invariant Feature Transform (SIFT)**
  -David Lowe's SIFT algorithm provided a robust method for detecting and describing local features in images, becoming essential in object recognition and matching.

- **2006: Histogram of Oriented Gradients (HOG)**
  -The HOG descriptor, introduced by Navneet Dalal and Bill Triggs, became a popular feature extraction technique, especially in human detection tasks.

KEY ADVANCEMENTS

2010-2014

The Deep Learning Revolution

- **2012: AlexNet and ImageNet**
  -Geoffrey Hinton, Alex Krizhevsky, and Ilya Sutskever introduced AlexNet, a deep convolutional neural network (CNN) that won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. This victory demonstrated the power of deep learning for computer vision tasks, revolutionizing the field.

- **2013: R-CNN (Region-based Convolutional Neural Networks)**
  -Ross Girshick introduced R-CNN, a model that significantly improved object detection performance by combining region proposals with CNNs, leading to more accurate detection and classification of objects within images.

KEY ADVANCEMENTS

2015-2019

Evolution of Applications and Improved Architectures

- **2014-2015: Generative Adversarial Networks (GANs)**
  - Ian Goodfellow introduced GANs, a class of machine learning frameworks that became widely used in image generation, manipulation, and enhancement. Furthering new possibilities in computer vision.

- **2015: ResNet (Residual Networks)**
  - Kaiming He and his team developed ResNet, a deep neural network architecture that solved the problem of vanishing gradients in very deep networks, enabling the training of much deeper CNNs. ResNet became foundational in tasks like image classification and object detection.

- **2015: Faster R-CNN**
  - Ross Girshick and collaborators introduced Faster R-CNN, an improvement over R-CNN that included a Region Proposal Network (RPN), significantly speeding up object detection processes.

- **2016: YOLO (You Only Look Once)**
  - Joseph Redmon introduced YOLO, an object detection model known for its speed and efficiency. Unlike previous methods, YOLO framed object detection as a single regression problem, allowing real-time object detection.

KEY ADVANCEMENTS

2020-Present

Integration of AI and Real-Time Applications

- **2020: Vision Transformers (ViT)**
  - The introduction of Vision Transformers by Google Research marked a shift from CNNs to transformer models for vision tasks. ViTs apply the transformer architecture, originally developed for NLP, to image recognition tasks, achieving state-of-the-art results.

- **2020: CLIP (Contrastive Language–Image Pretraining)**
  - Developed by OpenAI, CLIP demonstrated the ability to perform zero-shot learning on various image classification tasks by training on a wide variety of image-text pairs, greatly expanding the versatility of computer vision models.

- **2021: DALL-E and Image Synthesis**
  - OpenAI introduced DALL-E, a model capable of generating images from textual descriptions. This advancement in multimodal AI demonstrated significant progress in the ability to understand and generate visual content based on natural language prompts.

- **2023: Stable Diffusion**
  - The release of Stable Diffusion, a text-to-image generation model, brought high-quality, generative AI capabilities to the mainstream, furthering the integration of computer vision with creative and design industries.

FUTURE TRAJECTORY OF COMPUTER VISION

- The future of computer vision will definitely involve deeper integration of AI and robotics, more sophisticated autonomous systems, enhanced human computer interactions, and advanced predictive analytics. Areas such as self-supervised learning, explainable AI, and enhanced 3D vision will continue to drive evolve.

- As computer vision technology becomes more pervasive, ethical and privacy concerns will play an increasingly important role. Ensuring responsible use of computer vision, addressing biases in models, and protecting user privacy will be very difficult challenges in the future.

- Continued advancements will enable even more powerful real time vision applications in various domains, including autonomous vehicles, smart cities, and augmented reality.

THE ENID