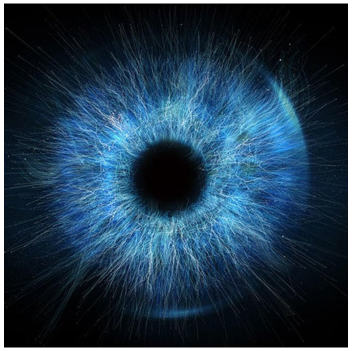
Computer Vision: Object: Classification, Identification, Verification, Landmark Detection, Segmentation, and Recognition.



Our Future: Tesla's self-driving cars. Microsoft Kinect Pose detection for medicine and entertainment. Medical imaging for detection of brain tumors and cancer.

Computer Vision: How It Is Changing Our Modern World.

Jason Ismail
Bellevue University
Bellevue, NE 68005, USA
jismail@my365.bellevue.edu

Shilpa Kolekar
Bellevue University
Bellevue, NE 68005, USA
skolekar@my365.bellevue.edu

Anna Harvey
Bellevue University
Bellevue, NE 68005, USA
anharvey@my365.bellevue.edu

Hedyeh Erfani
Bellevue University
Bellevue, NE 68005, USA
herfani@my365.bellevue.edu

Author Keywords

Computer Vision; Machine Learning; Clustering Models; Anomaly Detection; Prediction; and Association-rule Mining; Data Science; Self Driving Cars; Neuro Medicine; Microsoft Kinect Pose detection. Object: Classification, Identification, Verification, Landmark Detection, Segmentation, and Recognition.

ACM Classification Keywords

I.2 ARTIFICIAL INTELLIGENCE
I.2.6 Learning (K.3.2) Connectionism and neural nets
I.2.9 Robotics: Autonomous vehicles
I.2.10 Vision and Scene Understanding (I.4.8, I.5)
3D/stereo scene analysis, Architecture and control structures [**]
Intensity, color, photometry, and thresholding Modeling and recovery of physical attributes Motion Perceptual reasoning Representations, data structures, and transforms Shape Texture Video analysis

Abstract

Computer Vision: The growing field of computer vision lies at the heart of a multitude of various modern technologies. The amount of visual data collected continues to rise, making computer vision an essential tool for processing and organizing visual data. It incorporates several of the methods used by data scientists, including clustering, anomaly detection, association-rule mining, and prediction. This technology is a driving force across different domains. Major companies, including Google, Facebook, and Amazon, have invested heavily in this technology while inspiring innovations such as Tesla's self-driving cars, reCAPTCHA verification systems, and advances in medicine. In this project, we will conduct a general survey of the field of computer vision, explore how the technology works, assess its current challenges, and discuss how different industries put computer vision into action. Using computer vision, data scientists have an opportunity to optimize the flood of visual data to meet the needs of our global community.

Why is it important?

Visual Data in Abundance: In 2005, YouTube posted its first video. Since then, the amount of visual data has increased exponentially. In May of 2019, according to Statista.com, 500 hours of video were uploaded by users every minute. In 2019, the United States Census

Microsoft Kinect Pose Detection: Algorithms used in conjunction with Kinect can allow us to accurately model human movement.

Our Future: Pose detection to accurately diagnose and track ailments. Disease prognosis would be better understood, and life expectancy increased.

Bureau reported that there were 328.2 million people in the United States, of which 269.44 million of them have a smartphone capable of taking high-quality video. Cameras and hard drive space have become inexpensive, making it possible for companies to collect large quantities of data. If data scientists can use computers to interpret this new influx of data, it would open the doors to a multitude of different types of projects. Examples of this are already evident in our modern society. Visual data is currently used in the medical industry to discover and classify brain tumors and help diagnose patients with multiple sclerosis. People are identified and tracked to recognize possible threats to security systems. Shopping can become quicker and more efficient. Roads could become safer with self-driving cars.

Computer vision: solves the following problems.

Object Classification: This focus allows the computer to discover what general object category exists within an image. It finds objects that the network has been trained to see. Examples include cars, people, crosswalks, stop signs, cats, and dogs.

Object Identification: Once it classifies what kind of object it has found, it tries to classify the object with more detail. Examples include identifying car models: Ford Explorer, Toyota Corolla, Chevrolet Impala. Dogs are classified by breed, such as Labrador retriever, German shepherd, or French bulldog and may include identifying aspects like color.

Object Verification: This process finds out if a specific object is in the image. Examples include whether or not a person, car, or cat is in the image.

Object Detection: Locations of objects can be identified within the image. Object detection can be useful for tracking objects like humans and cars.

Object Landmark Detection: Objects can be broken down even further. Examples include focusing on a dogs' eyes, nose, and mouth, a brain tumor outline, or finding the car door.

Object Segmentation: This defines which pixels belong to specific objects. It is a way of separating subjects from other items within the image.

Object Recognition: Multiple types of objects can be deciphered within an image. An excellent example of this is Tesla's self-driving cars which discover a multitude of items within an image, including people, crosswalks, roads, other vehicles, and stop signs.

Computer vision: is fundamentally different from human vision.

The Details: Working with visual data is fundamentally complex for a computer. To a computer, an image is a collection of pixels with numerical color values in the range of 0 and 255 for the primary colors red, green, and blue. It also makes distinctions between black and white. Because of this, unlike humans, computers have a tough time picking out what is essential within an image. If a person shows a 5-year-old a picture of a dog, they can quickly look at the picture of a dog and ignore everything else about the picture to decide that

Tesla Self Driving Cars: Computer Vision played significant role in Tesla's success while making self-driving cars.

Our Future: Self driving cars using computer vision, identify surrounding objects, which helps drivers by increasing level of automation and keeps our roads safer.

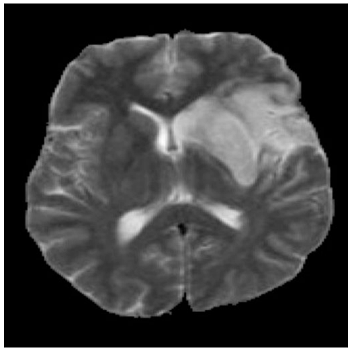
it is a dog. A computer, on the other hand, needs to look at a picture sequentially, pixel by pixel, to discover the subjects within the picture. It is not good enough to know just what color the pixels are, but the computer needs to make decisions based on patterns that the pixels are making within the image. For a computer to work with images, the computer needs to resize the image, so there are fewer pixels to evaluate. Then a computer scans the image to gather information about the image. It looks for groups of pixels that are found by interpreting edges within the image and uses this information to classify objects within the image. Sometimes to achieve this, it alters the image and then looks at the results. Then it looks at features found within each object to try and classify the object by using a neural network.

A neural network works by feeding a computer images of a specific subject type, for example cats and dogs. These images serve as the inputs for the neural network. The network takes the image of a cat or a dog and analyses it. Then it makes an educated guess on whether or not it is a cat or a dog. The system learns based on positive and negative feedback and builds variables for the features it found within the images. These variables serve as the hidden layers within the system. Each image is processed, and the computer continues to make guesses based on weights and probability within the nodes of its network. Then, based on positive or negative reinforcement, it goes back within the hidden layers and adjusts the weights assigned to the network based on the results. Eventually, once the network is functioning at a high level, there is no need to tell the computer whether or not it is looking at a cat or dog; it can discover it on its own.

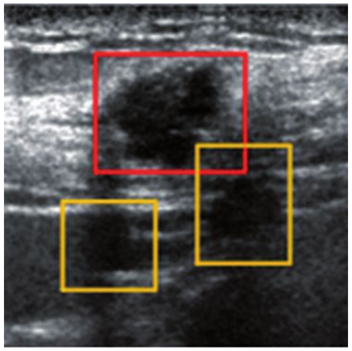
Self Driving Cars

Vehicle Automation: Tesla, a pioneer in the self-driving auto industry, has enabled computer vision, artificial intelligence, pattern recognition, and many other technologies into the self-driving car. Computer vision in these vehicles works by capturing visual data in real-time through cameras and takes footage from its surroundings. It then uses the data to create live 3-D image maps. The vehicle can understand its surroundings by collecting data from these images. If there is an obstacle, it makes changes in its route. The car uses the same 3-D map image to make predictions about the objects on the road, giving it the ability to see traffic lights, other surrounding vehicles, and pedestrians in real-time. Computer vision has enabled self-driving vehicles to identify obstacles and avoid accidents making our roads safer.

Vehicles can be classified into six levels of automation. Level 0: vehicles do not have any automation at all. Level 1: vehicles have some low-level features for driver assistance. Level 2: vehicles have the most autonomous features but still need drivers to pay full attention to the road. Level 3: vehicles are proper autonomous vehicles where the drivers do not need to pay full attention to the road, but they need to be ready to take over control when needed. Level 4: vehicles are autonomous vehicles that technically do not need a driver, but there are geographical restrictions, so a driver is still required. Level 5: vehicles are autonomous vehicles that do not need any drivers, and there are no restrictions. Tesla is currently at Level 2, and computer vision will be a significant contributing factor to take Tesla to Level 3 and beyond. Tesla is at Level 2 is because of certain limitations. For



Medical Imaging: Computer vision pre-processing strips the skull from the image to help focus on brain tissue to find stroke lesions.



Our Future: The red box shows a true tumor. The yellow boxes show a false tumor. Computer vision helps us distinguish images that are similar to the human eye.

example, if the lines on the road are faded, or the lines do not merge properly, computer vision will make mistakes in capturing visual data correctly. It would not be able to make appropriate decisions to navigate the vehicles. While Tesla works on identifying solutions for such limitations, they have already added several features that push towards making the cars autonomous. Some of the features are Traffic-Aware Cruise Control, Autosteer, Auto Lane Change, Auto Park, Summon, and Traffic and Stop Sign Control. As Tesla strides towards higher levels of automation, computer vision technology is going to play a significant role in their success.

Medical Imaging: Tumor Detection

Medical Industry: Medical imaging, such as computed tomography (CT) and magnetic resonance imaging (MRI), capture images within the body for medical analysis. These images can identify tumors, hemorrhages, clots, and other abnormalities. Many of these medical conditions can rapidly progress to a life-threatening stage. Medical professionals must analyze these images quickly to save lives. The application of computer vision to image analysis has aided doctors in identifying abnormalities more efficiently.

For abnormality identification, the first step of pre-processing involves removing the parts of the image that do not apply to the diagnosis. For example, when scanning for brain tumors or stroke lesions, they remove the skull structure from the original image. Pre-processing uses various algorithms to discern non-relevant parts of the image from the rest. Once completed, convolutional neural networks assist in image segmentation for image detection, edge detection, or filtering to distinguish the abnormalities from normal segments of the image. The network learns to identify abnormalities through supervised learning iterations. Probability calculations for medical risk of the abnormality may also be conducted at this stage through additional neural levels in the network. Algorithm and neural network customization are required between different medical areas, since each type of abnormality requires its own set of parameters for identification.

Computer vision innovation has progressed throughout the medical profession in recent years. In many cases, computer vision assisted analysis can accelerate the identification of abnormalities, be repeated with high sensitivity and high specificity, and be more accurate than professional visual analysis alone. It is a field that continues to grow and holds much promise for medical professionals and patients in the future.

Microsoft Kinect: Pose Detection

Entertainment & Medical: Microsoft Kinect Pose detection is a computer vision technique that predicts body poses by detecting figures and their movements. A computer model is used to track various joints efficiently. It successfully detects poses by taking an image and breaking it down to evaluate which joints and body parts are moving. With computer vision, the computer processes images frame by frame to understand what pose the individual is making. It then uses a camera to capture images at a resolution of 640x480 pixels at 30 frames per second. Kinect employs a depth camera, which works in conditions such as low lighting and then categorizes the images based on the poses the subject makes. It can handle a

Conclusion: Computer vision is an important technology. Training computers to be able to interpret image data will change our future.

References: Thank you for all of the research that has already been done in the field of computer vision and deep learning.

full range of body shapes and motions and categorizes them. The Kinect is used for gaming as well as the detection of multiple sclerosis.

Since multiple sclerosis causes physical impairment over an unpredictable timeframe, it is imperative to find ways to detect and assess the condition. Kinect can assist with pose detection; thus, it can also be used to detect abnormalities in motor function. One of the prototypes to do this is called ASSESS MS. With the help of Kinect and health professionals, in-depth videos are taken of the movement of patients. Then, the videos are given a classification based on motor ability per the Expanded Disability Severity Score (EDSS). Over time, these videos and the respective scores can be compared to analyze the progression of multiple sclerosis.

Conclusion

Computer vision's wide array of applications in multiple domains makes it a powerful tool for future development. It can be used to save lives, make lives easier, and help research progress in all fields. This complex technology will continue to develop. Data scientists are needed for further study in this field and to find more real-world applications. Computer vision has only begun to change our world.

Acknowledgements

Our sincere thanks to:

- Team members for excellent teamwork.
- Professor Fadi Alsaleem for providing constructive feedback and support.
- Computer Vision research community for their inspiration.

Academic References:

1. Anandhavalli Muniasamy and Areej Alasiry. 2020. Deep Learning: The Impact on Future eLearning. *International Journal of Emerging Technologies in Learning (iJET)* 15, 01 (2020), 188. DOI:http://dx.doi.org/10.3991/ijet.v15i01.11435

2. Cecily Morrison et al. 2016. Assessing Multiple Sclerosis With Kinect: Designing Computer Vision Systems for Real-World Use. *Human–Computer Interaction* 31, 3-4 (2016), 191–226. DOI:http://dx.doi.org/10.1080/07370024.2015.1093421

3. Chao Tao et al. 2019. New one-step model of breast tumor locating based on deep learning. *Journal of X-Ray Science and Technology* 27, 5 (2019), 839–856. DOI:http://dx.doi.org/10.3233/xst-190548

4. Chao Zhou et al. 2019. Method for segmentation of overlapping fish images in aquaculture. *Method for segmentation of overlapping fish images in aquaculture* 12, 6 (2019), 135–142. DOI:http://dx.doi.org/10.25165/j.ijabe.20191206.3217

5. Chaohui Tang, Qingxin Zhu, Wenjun Wu, Wenlin Huang, Chaoqun Hong, and Xinzheng Niu. 2020. PLANET: Improved Convolutional Neural Networks with Image Enhancement for

Image Classification. *Mathematical Problems in Engineering* 2020 (March 2020), 1–10. DOI:http://dx.doi.org/10.1155/2020/1245924

6. Jelena Kocić, Nenad Jovičić, and Vujo Drndarević. 2019. An End-to-End Deep Neural Network for Autonomous Driving Designed for Embedded Automotive Platforms. *Sensors* 19, 9 (March 2019), 2064. DOI:http://dx.doi.org/10.3390/s19092064

7. Jeng-Ywan Jeng, Tzuoh-Fei Mau, and Sh Yeu-Ming Leu. 2000. Computer integrated vision and neural network technologies for intelligent automatic laser butt joint welding. *International Journal of Computer Integrated Manufacturing* 13, 4 (2000), 345–357. DOI:http://dx.doi.org/10.1080/095119200407705

8. Muhammad Aqeel Aslam and Daxiang Cui. 2017. Brain Tumor Detection from Medical Images: A Survey. *Nano Biomedicine and Engineering* 9, 1 (2017). DOI:http://dx.doi.org/10.5101/nbe.v9i1.p72-81

9. Nikhil Naik, Ramesh Raskar, and César A. Hidalgo. 2016. Cities Are Physical Too: Using Computer Vision to Measure the Quality and Impact of Urban Appearance. *American Economic Review* 106, 5 (2016), 128–132. DOI:http://dx.doi.org/10.1257/aer.p20161030

10. Paulo J.S. Gonçalves, Bernardo Lourenço, Samuel Santos, Rodolphe Barlogis, and Alexandre Misson. 2020. Computer Vision Intelligent Approaches to Extract Human Pose and Its Activity from Image Sequences. *Electronics* 9, 1 (2020), 159. DOI:http://dx.doi.org/10.3390/electronics9010159

11. R. Karthik and R. Menaka. 2017. Computer-aided detection and characterization of stroke lesion – a short review on the current state-of-the art methods. *The Imaging Science Journal* 66, 1 (August 2017), 1–22. DOI:http://dx.doi.org/10.1080/13682199.2017.1370879

12. Xuguo Yan, Long Wen, and Liang Gao. 2019. A Fast and Effective Image Preprocessing Method for Hot Round Steel Surface. *Mathematical Problems in Engineering* 2019 (2019), 1–14. DOI:http://dx.doi.org/10.1155/2019/9457826

13. Yasunari Miyagi, Kazuhiro Takehara, and Takahito Miyake. 2019. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images. *Molecular and Clinical Oncology* (April 2019). DOI:http://dx.doi.org/10.3892/mco.2019.1932

Non-Academic

14. Anon.Retrieved May 14, 2020 from https://www.cs.ryerson.ca/~aharley/vis/conv/flat.html

15. Anon. 2020. Autopilot and Full Self-Driving Capability. (April 2020). Retrieved May 14, 2020 from https://www.tesla.com/support/autopilot

16. Anon. Assessing Multiple Sclerosis with Kinect. Retrieved May 14, 2020 from https://www.microsoft.com/en-us/research/project/assessing-multiple-sclerosis-with-kinect/

17. Anon. Brain Tumor Detection using Mask R-CNN. Retrieved May 14, 2020 from https://www.kdnuggets.com/2020/03/brain-tumor-detection-mask-r-cnn.html

18. Anon. The Future Now - A Fighter Pilot's View On Tesla Autopilot. Retrieved May 14, 2020 from https://www.torquenews.com/6968/future-now-fighter-pilot-s-view-tesla-autopilot

19. Anon. Kaggle. Retrieved May 14, 2020 from https://www.facebook.com/kaggle/photos/a.10150387148668464/10157044315513464/?type=3Anon. An Overview of Human Pose Estimation with Deep Learning. Retrieved May 14, 2020 from https://www.kdnuggets.com/2019/06/human-pose-estimation-deep-learning.html

20. Elise Andrew. 2019. Google's AI Can Dream, and Here's What it Looks Like. (March 2019). Retrieved May 14, 2020 from https://www.iflscience.com/technology/artificial-intelligence-dreams/

21. freeCodeCamp.org. 2018. An intuitive guide to Convolutional Neural Networks. (February 2018). Retrieved May 14, 2020 from https://www.freecodecamp.org/news/an-intuitive-guide-to-convolutional-neural-networks-260c2de0a050/

22. Henok Tilaye and Haylemicheal Berihun. 2019. Image classification with convolutional neural networks. (August 2019). Retrieved May 14, 2020 from https://medium.com/analytics-vidhya/image-classification-with-convolutional-neural-networks-ac14a978f0fa

23. Ilija Mihajlovic. 2020. Everything You Ever Wanted To Know About Computer Vision. Here's A Look Why It's So Awesome. (February 2020). Retrieved May 14, 2020 from https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e

24. J. Clement. 2019. YouTube: hours of video uploaded every minute 2019. (August 2019). Retrieved May 14, 2020 from https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/

25. James Le. 2020. The 5 Computer Vision Techniques That Will Change How You See The World. (January 2020). Retrieved May 14, 2020 from https://heartbeat.fritz.ai/the-5-

computer-vision-techniques-that-will-change-how-you-see-the-world-1ee19334354b

26. James O'Malley. 2018. Captcha if you can: how you've been training AI for years without realising it. (January 2018). Retrieved May 14, 2020 from https://www.techradar.com/news/captcha-if-you-can-how-youve-been-training-ai-for-years-without-realising-it

27. Jamie Shotton et al. 2011. Real-time human pose recognition in parts from single depth images. *Cvpr 2011* (2011). DOI:http://dx.doi.org/10.1109/cvpr.2011.5995316

28. Jason Murdock. 2020. Artificial intelligence that can evolve on its own is being tested by Google scientists. (April 2020). Retrieved May 14, 2020 from https://www.newsweek.com/google-brain-artificial-intelligence-machine-learning-research-evolution-automl-zero-1497781?utm_source=ONTRAPORT-email-broadcast&utm_medium=ONTRAPORT-email-broadcast&utm_term=&utm_content=Data+Science+Insider:+April+17th,+2020&utm_campaign=18042020

29. Jennifer Whitlock and Msn. 2020. The History of Surgery: A Timeline of Medicine. (April 2020). Retrieved May 14, 2020 from https://www.verywellhealth.com/the-history-of-surgery-timeline-3157332

30. Jianfeng Zhao, Bodong Liang, and Qiuxia Chen. 2018. The key technology toward the self-driving car. *International Journal of Intelligent Unmanned Systems* 6, 1 (February 2018), 2–20. DOI:http://dx.doi.org/10.1108/ijius-08-2017-0008

31. Jill Goldsmith. 2020. Bill Gates Exits Microsoft Board To Focus On Philanthropy Full-Time, Ends An Era At The Company. (March 2020). Retrieved May 14, 2020 from https://deadline.com/2020/03/bill-gates-microsoft-board-philanthropy-1202882963/

32. Josh Patterson and Kirit Basu. 2017. Solving real-world business problems with computer vision. (September 2017). Retrieved May 14, 2020 from https://www.oreilly.com/radar/solving-real-world-business-problems-with-computer-vision/

33. Matthew Beedham. 2020. The 6 levels of autonomous driving, explained as fast as possible. (February 2020). Retrieved May 14, 2020 from https://thenextweb.com/shift/2020/02/27/6-levels-autonomous-self-driving-explained-tesla-waymo-autopilot/

34. Myroslava Zelenska. 2019. Neural Networks-Relation to Human Brain and Cognition. (August 2019). Retrieved May 14, 2020 from https://becominghuman.ai/neural-networks-

relation-to-human-brain-and-cognition-b45575359f64

35. Nahua Kang. 2019. Multi-Layer Neural Networks with Sigmoid Function- Deep Learning for Rookies (2). (February 2019). Retrieved May 14, 2020 from https://towardsdatascience.com/multi-layer-neural-networks-with-sigmoid-function-deep-learning-for-rookies-2-bf464f09eb7f

36. Natalie Fletcher. How facial recognition and the right AI provider can improve your operations. Retrieved May 14, 2020 from https://www.clarifai.com/blog/how-the-right-facial-recognition-provider-can-improve-your-operations

37. Sima. Quality Eyeglass Lenses in Southampton, NY. Retrieved May 14, 2020 from https://www.hamptonsvision.com/eyeglasses/quality-eyeglass-lenses-in-southampton-ny/

38. Stéphane D'Angelo. 2017. Test Just Dance 2017 Nintendo Switch : bouge ton joy-con ! (March 2017). Retrieved May 14, 2020 from https://www.thmmagazine.fr/jeux-video/test-just-dance-2017-nintendo-switch/

39. ThingLink. Human Brain by Alyssa Burnham. Retrieved May 14, 2020 from https://www.thinglink.com/scene/826181179861368833

40. Vibhor Nigam. 2020. Understanding Neural Networks. From neuron to RNN, CNN, and Deep Learning. (February 2020). Retrieved May 14, 2020 from https://towardsdatascience.com/understanding-neural-networks-from-neuron-to-rnn-cnn-and-deep-learning-cd88e90e0a90

41. WorldQuant. 2019. Understanding Images: Computer Vision in Flux. (January 2019). Retrieved May 14, 2020 from https://www.weareworldquant.com/en/thought-leadership/understanding-images-computer-vision-in-flux/