

Data Science with Images: Computer Vision

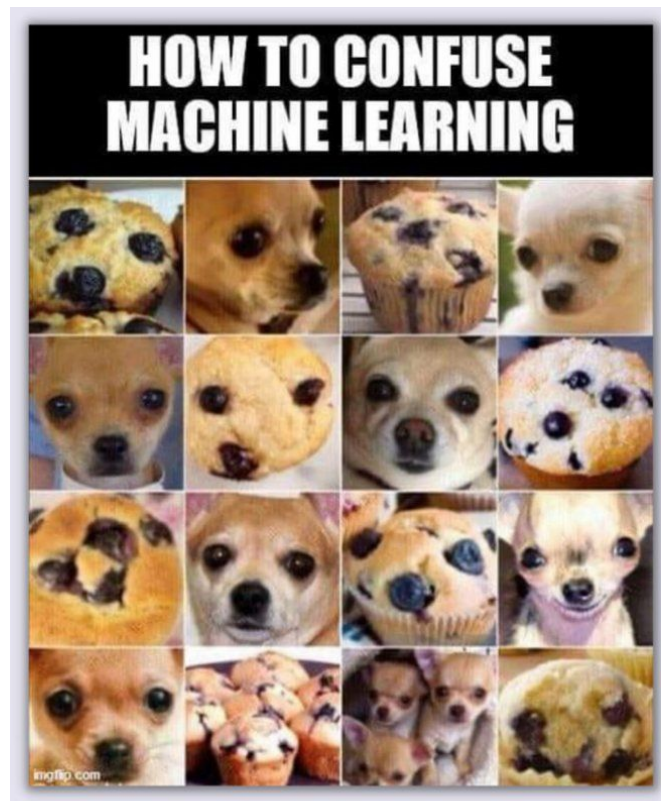


Week 11 Session 1

Lecture and Slides by Jazzie R. Jao | jazzie.jao@dlsu.edu.ph

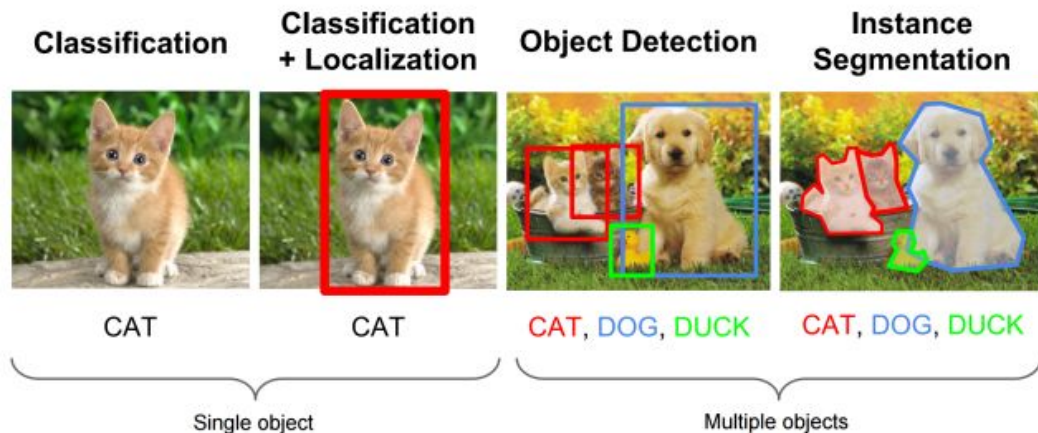
What is Computer Vision

- Computer vision is a field within data science focused on enabling computers to interpret and analyze visual information.
- We can develop models that "see" and interpret images, detect patterns, recognize objects, and even make decisions based on visual inputs.
- Computer vision is implemented in Python using popular libraries like
 - OpenCV, NumPy, and PyTorch



Basic Principles of Computer Vision in Python

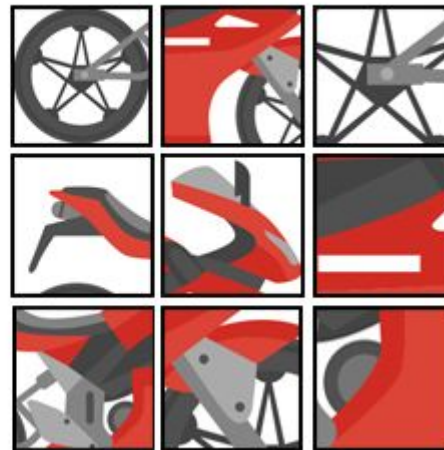
- Computer vision involves tasks such as image classification, object detection, segmentation, and feature extraction.



Basic Principles of Computer Vision in Python

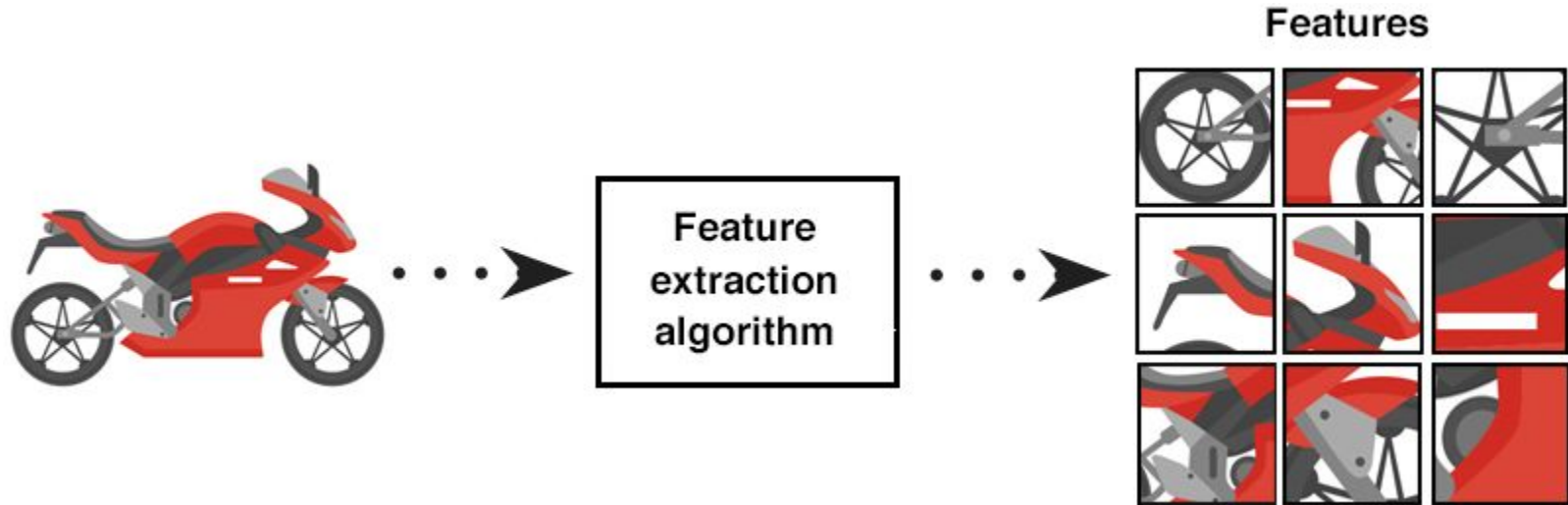
- Computer vision involves tasks such as image classification, object detection, segmentation, and feature extraction.

Features



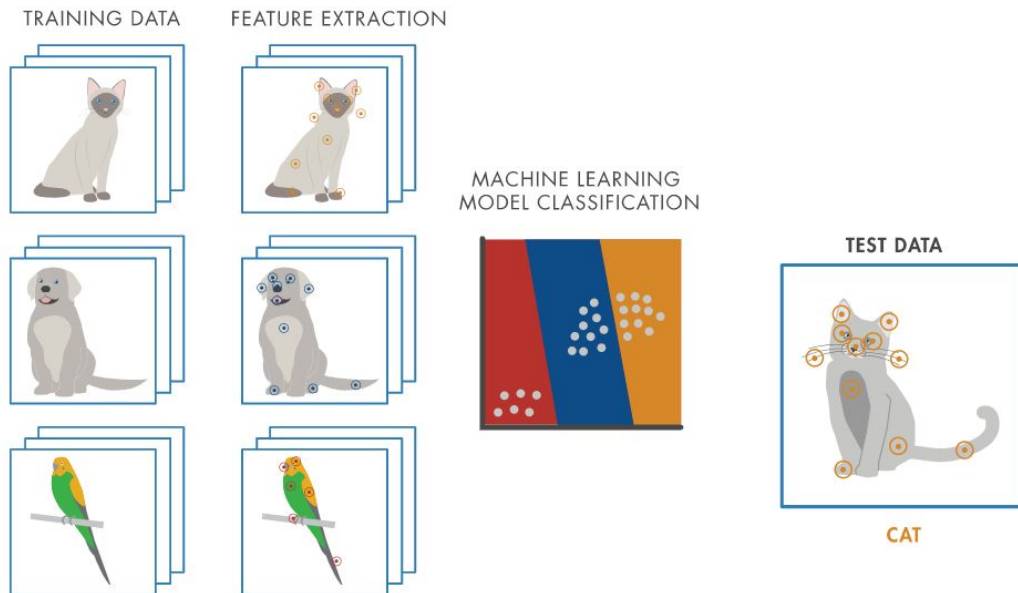
Basic Principles of Computer Vision in Python

- Computer vision involves tasks such as image classification, object detection, segmentation, and feature extraction.



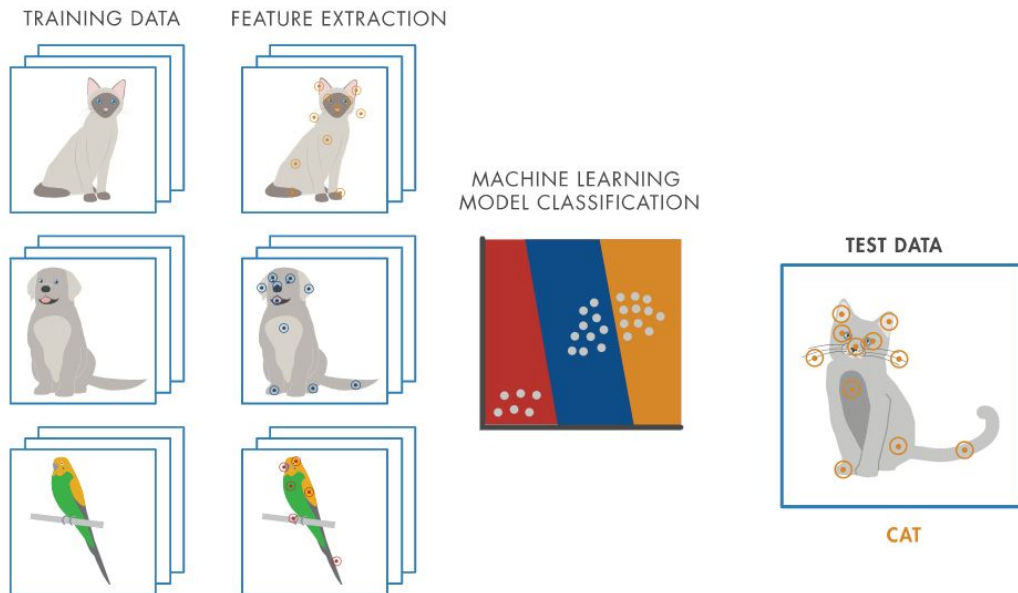
Basic Principles of Computer Vision in Python

- Computer vision involves tasks such as image classification, object detection, segmentation, and feature extraction.



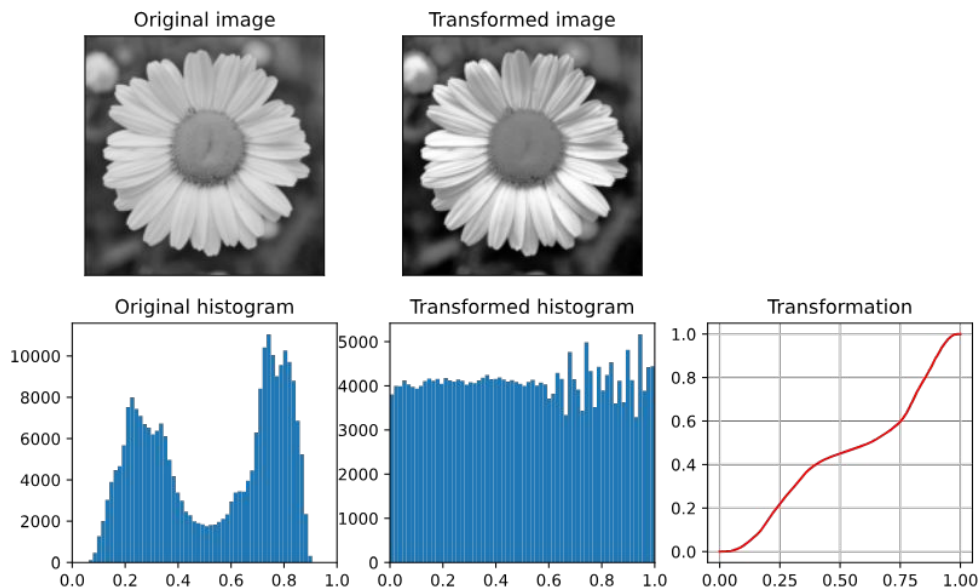
Basic Principles of Computer Vision in Python

- Computer vision involves tasks such as image classification, object detection, segmentation, and feature extraction.



Basic Principles of Computer Vision in Python

- At its core, it translates visual data into meaningful information using mathematical models, often driven by machine learning and deep learning techniques. Some foundational principles include
 - **Image Processing:** This involves manipulating image data (e.g., adjusting brightness, resizing, filtering) to prepare it for analysis.




Basic Principles of Computer Vision in Python

- At its core, it translates visual data into meaningful information using mathematical models, often driven by machine learning and deep learning techniques. Some foundational principles include
 - **Feature Extraction:** Detecting significant features (like edges or textures) in images, which can then be used for pattern recognition.

 You



can you please describe this image in details?

 ChatGPT

The image is a playful visual composition that juxtaposes blueberry muffins with the faces of Chihuahuas to highlight their similarities. There are eight close-up pictures arranged in two rows, with each row containing four images. The images alternate between blueberry muffins and the faces of Chihuahuas. The muffins have blueberries that are positioned where the eyes and nose would be on a face, creating an amusing resemblance to the Chihuahuas' facial features. The Chihuahuas have expressive eyes and noses that mimic the appearance of the blueberries on the muffins. This visual pun is commonly shared on social media to entertain viewers with the striking resemblance between the two, playing on the notion that sometimes everyday objects can unexpectedly mimic life.

 You

what is the third image on the top row?

 ChatGPT

The third image on the top row is a muffin. It can be identified by the baked, crumbly texture typical of a muffin and the blueberries that look like eyes and a nose.

Basic Principles of Computer Vision in Python

- At its core, it translates visual data into meaningful information using mathematical models, often driven by machine learning and deep learning techniques. Some foundational principles include
 - **Machine Learning:** Training algorithms on datasets of images to recognize specific patterns, objects, or actions.

You



can you please describe this image in details?

ChatGPT

The image is a playful visual composition that juxtaposes blueberry muffins with the faces of Chihuahuas to highlight their similarities. There are eight close-up pictures arranged in two rows, with each row containing four images. The images alternate between blueberry muffins and the faces of Chihuahuas. The muffins have blueberries that are positioned where the eyes and nose would be on a face, creating an amusing resemblance to the Chihuahuas' facial features. The Chihuahuas have expressive eyes and noses that mimic the appearance of the blueberries on the muffins. This visual pun is commonly shared on social media to entertain viewers with the striking resemblance between the two, playing on the notion that sometimes everyday objects can unexpectedly mimic life.

You

what is the third image on the top row?

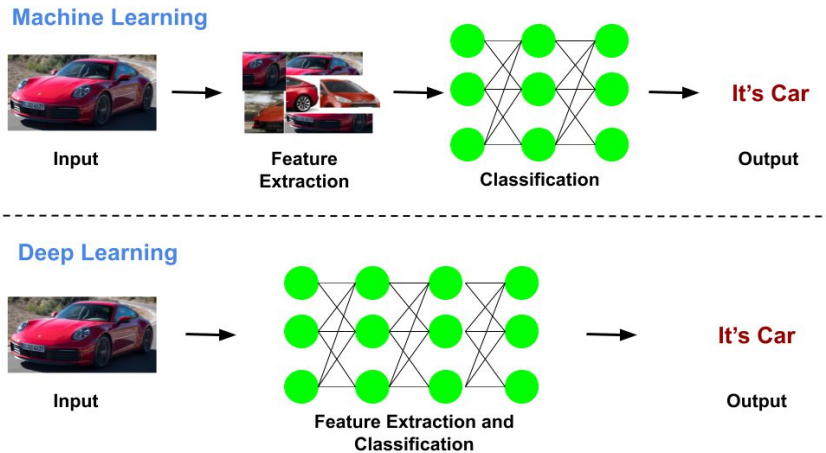
ChatGPT

The third image on the top row is a muffin. It can be identified by the baked, crumbly texture typical of a muffin and the blueberries that look like eyes and a nose.



Basic Principles of Computer Vision in Python

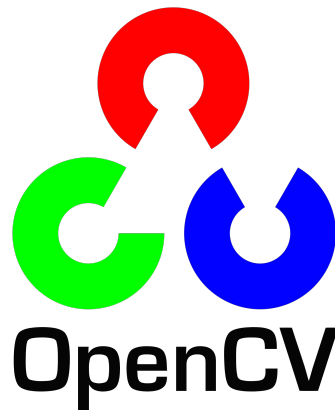
- At its core, it translates visual data into meaningful information using mathematical models, often driven by machine learning and deep learning techniques. Some foundational principles include
 - **Machine Learning:** Training algorithms on datasets of images to recognize specific patterns, objects, or actions.
 - **Deep Learning:** Using convolutional neural networks (CNNs) to automatically learn complex visual features and hierarchies from large image datasets.



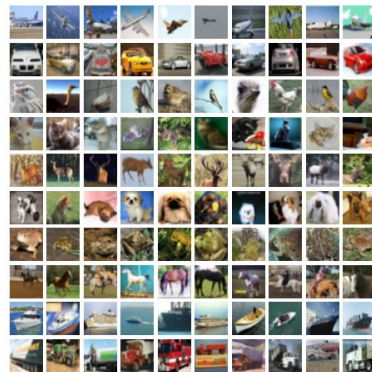
Toolbox for Computer Vision in Python

To work with computer vision in Python, you'll need the following libraries:

- **OpenCV:** Provides many functions for image processing and feature extraction.
- **NumPy:** Useful for numerical operations and image manipulation.
- **PyTorch or TensorFlow:** Popular deep learning frameworks for building and training neural networks.
- **Pre-trained Models and Datasets:** Using pre-trained models from packages like torchvision (e.g., ImageNet) is common for applications like object detection and classification.



airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck

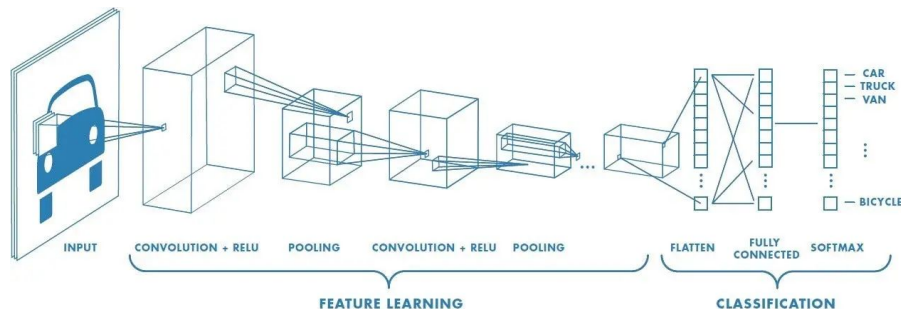
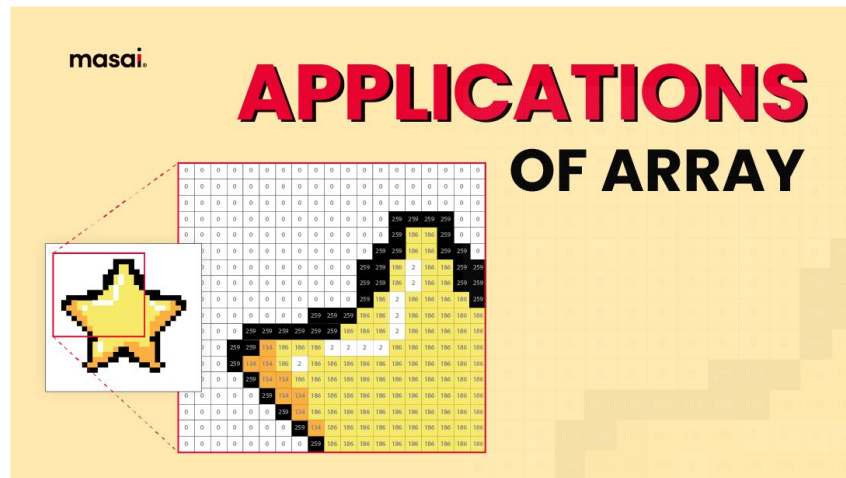


VS PYTORCH

Mathematical Foundations

The key to understanding computer vision lies in the mathematical foundations of image data and transformations:

- **Images as Arrays:** An image is represented as a matrix where each pixel has a value indicating color or intensity.
- **Convolutional Neural Networks (CNNs):** These use convolutional layers to process image data, where filters (kernels) slide across the image to detect features (edges, textures).

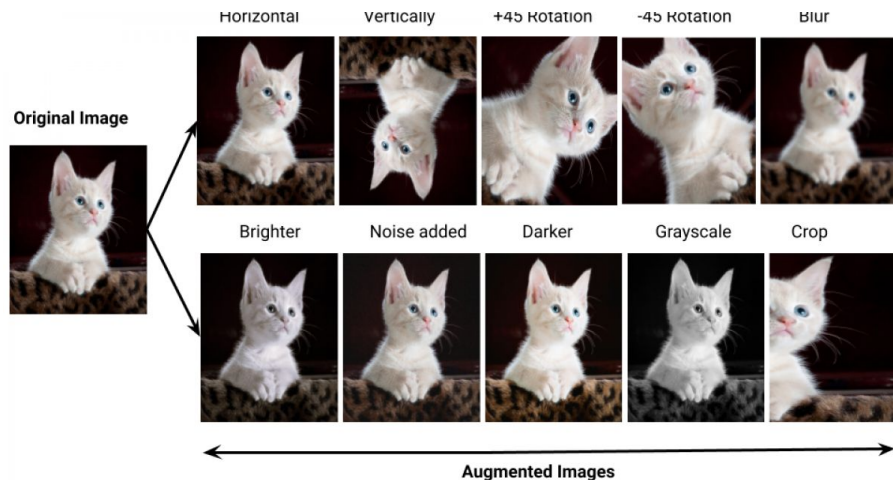
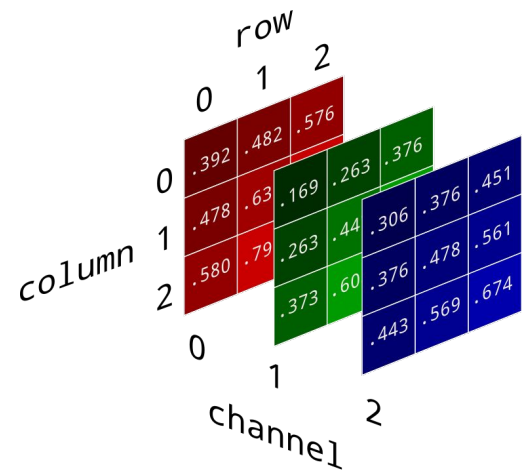


Representations

Matrix. An image is typically represented as a matrix of pixel values. In grayscale images, each pixel holds one intensity value (0 to 255), while in color images (RGB), each pixel has three values for red, green, and blue channels.

Normalization: Since pixel values range from 0 to 255, it's common to normalize them (e.g., between 0 and 1 or with a mean and standard deviation) to improve model training stability.

Resizing and Augmentation. Image data often varies in size, so resizing helps standardize input. Data augmentation techniques like rotation, flipping, and cropping introduce variations, making models more robust by simulating diverse conditions.



Representations

Matrix. An image is typically represented as a matrix of pixel values. In grayscale images, each pixel holds one intensity value (0 to 255), while in color images (RGB), each pixel has three values for red, green, and blue channels.

Normalization: Since pixel values range from 0 to 255, it's common to normalize them (e.g., between 0 and 1 or with a mean and standard deviation) to improve model training stability.

Resizing and Augmentation. Image data often varies in size, so resizing helps standardize input. Data augmentation techniques like rotation, flipping, and cropping introduce variations, making models more robust by simulating diverse conditions.

