13.11.2025

OPENCAMPUS.sh

Machine Learning With TensorFlow

CONVOLUTIONAL NEURAL NETWORKS IN TENSORFLOW – PART I

- Quiz
- Assignment
- Input
- (Open Discussion)
- Project Milestones

CNNs in TensorFlow Part I – Formular ausfüllen

QUIZ

https://forms.office.com/e/A05G75dgNm

ASSIGNMENT

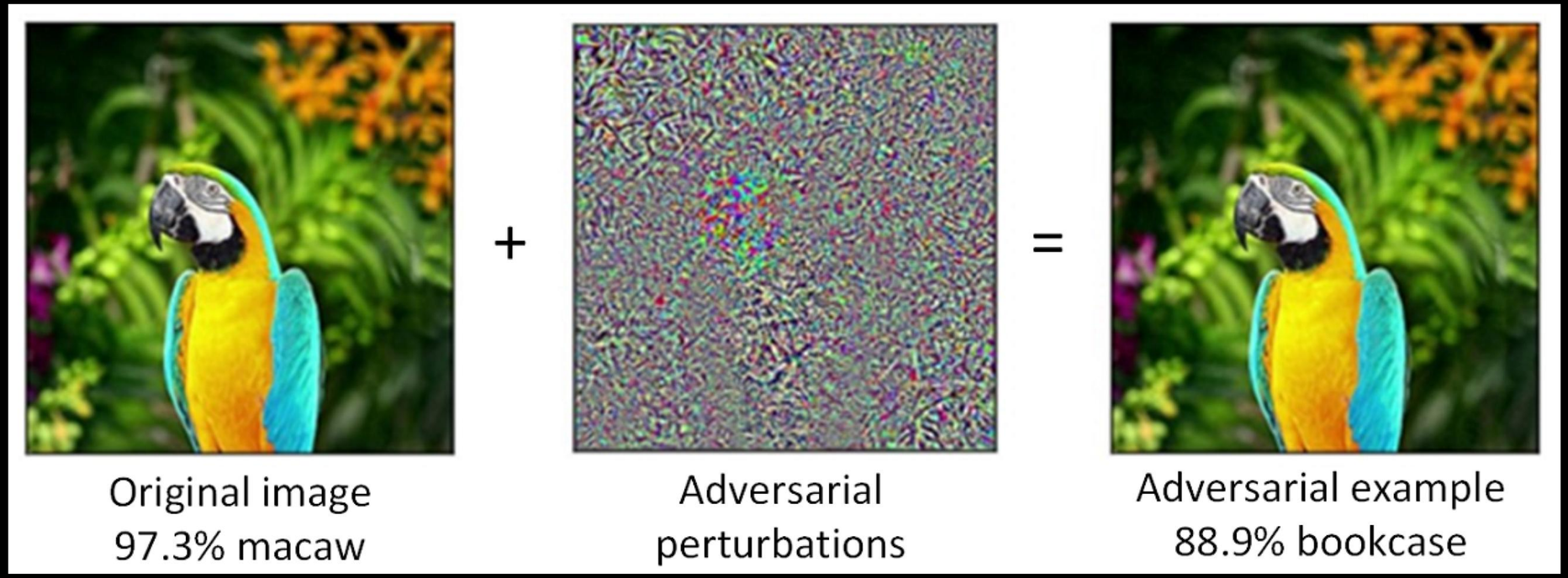
ASSIGNMENT: WHO WILL PRESENT NEXT?

KEY CNN CONCEPTS (IN COMPUTER VISION)

- **Kernel**: a matrix of weights applied across the input image (slid over the image: only ever applied to a small part of the image at a time)
- **Stride**: number of pixels by which the kernel moves across the image
- **Padding**: extra pixels (usually zeros, i.e., black) added around the image so the kernel can slide over the edges of the image
- **Feature map**: 2D output from applying one filter over the image
- **Pooling**: reduces information from multiple pixels to a single value

ADVERSARIAL ATTACKS

- Neural Networks do not „see“ the same way that humans do
- See adversarial attacks (Shi et al., 2020):



PROJECTS MILESTONES

30.10.	Present your Ideas
06.11.	Form Groups
13.11.	Literature Review
20.11.	Dataset Characteristics

Deadline for completing the repo sections: 23.11.

27.11.	Individual Feedback Sessions
04.12.	Baseline Model Estimation
11.12.	Definition of Model Evaluation

Deadline for completing the repo sections: 14.12.

18.12.	Individual Feedback Sessions
15.01.	Project Presentations, Part I
22.01.	Project Presentations, Part II

Submission deadline for the documented repo: 30.02.

NEXT MILESTONE: DATASET CHARACTERISTICS

- Perform a data analysis of your dataset to identify:
  - Key characteristics of the dataset (Original data sources, number of samples, etc.)
  - Potential biases arising from the collected data
  - Outliers in the dataset
  - Missing values
  - Feature distributions
  - Correlations

- Utilize graphics, plots and tables to summarize the main characteristics of the dataset.

DATASET CHARACTERISTICS

DATASET CHARACTERISTICS

- Write down key points on the most important aspects of how your data was collected
- What are potential biases?
- Are there outliers in the dataset?
- For classification tasks: Are the classes balanced?
- Are there potential data augmentation approaches you can use?

FIRST DATA INSPECTION

- df.head(), df.tail():
  Display the first/last rows
- df.sample():
  Display a random sample of rows
- df.shape:
  Gives the number of rows and columns
- df.info():
  Summary of the data frame (including missing values, data types)

DESCRIPTIVE STATISTICS

- df.describe():
  Statistical summary of the numerical columns

- df.isnull():
  Checks for missing values (NULL / NA / nan)

VISUALIZATIONS

- Scatter plots
- Box plots
- Bar charts
- Histograms
- …

OPEN DISCUSSION

- Beside image data augmentation through transformations of the training data presented in the videos, what other forms of extending the diversity of an existing image dataset can you think of?

- What are the main differences between extending a dataset through (synthetically) generated images and performing image data augmentation during the training process?

TASKS UNTIL NEXT WEEK

- **Analyze the characteristics your dataset and add the analysis to your GitHub repository (Deadline: 23.11. 23:59**

- **Complete the learning material from the course handbook**

- **Complete assignment notebook from the course handbook**