Group Project

ARI 2129 - Principles of Computer Vision for AI

Deadlines:

Part A: 8th May 2024
Part B: 9th June 2024

Group Size: **Maximum of 4 students (Equal effort will be assumed)**

Maximum Marks: **100 (50% of the entire study unit)**

Plagiarism and Authenticity:

- No plagiarism will be tolerated. All sources must be adequately referenced in the code and documentation.
- Generative AI (GenAI) is allowed as long it is documented and in line with the instructions in this documentation and reported in the Generative AI Journal
- The use of generative AI needs to be declared and documented.
- Students may be asked for impromptu interviews to check their knowledge of the submitted content. Any information in the submitted documentation not being adequately explained by students and without its respective sources will be immediately reported as plagiarism.

Marking Scheme:

- 20% for documentation (maximum of 4 pages for Part 1 and 5 pages for Part 2)
  - 10% for documentation
  - 10% for Generative AI Journal (See guidelines below)
- 40% for Part A
  - 30% for Question 1 (Background)
  - 10% for Question 2 (Using GenAI as a writing assistant)
- 40% for Part B
  - 20% for Question 3 (Data Augmentation in Tensorflow and Pytorch)
  - 20% for Question 4 (Implementation of Data Augmentation in OpenCV)

Deliverables:

- One PDF file for each part is to be submitted on time according to the deadlines provided above.
- One PDF for the Generative AI Journal
- The documents should have a maximum of 4 pages for Part A, and another maximum of 5 pages for Part B. Questions with their title are to be set as headings in the document, clearly indicating which ones were not attempted, if any.
- Python source code in a Jupyter Notebook for each question as applicable.
- Each submission should be presented as a single Zip file with the Surname and Name of the student and the subject code, for example, SeychellDylanARI2129.zip.

    - In the zip file, the documentation must be in PDF and a folder containing the notebooks with any images needed.
    - In group submissions, a single submission is enough, and the names must be visible on the title page of the documentation.
    - The notebooks need to work as extracted from the zip file without the need for further configuration. This means that any images library will be accessed within the folders themselves or via URL.
- Only one submission per group is required, and the team members' names must be clearly outlined on the first page of the documentation.

Preliminaries

Project Aim and Objectives

This project explores data augmentation techniques in computer vision, including theoretical and practical aspects. It is based on the following objectives:

- To investigate the concept of data augmentation in computer vision, including its history, applications, and impact on deep learning.
- To analyse data augmentation in computer vision by reviewing the literature on the topic.
- To implement data augmentation techniques in mainstream frameworks such as TensorFlow and PyTorch and from first principles using OpenCV.
- To generate augmented images using the selected techniques and evaluate their output compared with other approaches and the original images.

Background

Data augmentation is a technique used in machine learning to artificially increase the size and diversity of training data by creating new examples similar to the original data. Data augmentation can be achieved through various transformations, including photometric distortions and geometric distortions.

Photometric distortions alter the image's appearance without changing its geometric properties. These transformations include changes in brightness, contrast, colour, saturation, and adding noise or blur to the image. By applying these transformations to the training data, data augmentation can help improve the robustness of the model to variations in lighting and image quality.

Geometric distortions are transformations that alter the geometric properties of the image, such as its size, orientation, and shape. These transformations include rotations, translations, scaling, and flipping. By applying these transformations to the training data,

data augmentation can help improve the model's robustness to variations in object position, pose, and scale.

In computer vision, data augmentation is particularly important because it can help overcome the limitations of limited training data, reduce overfitting, and improve the generalisation performance of deep learning models. The next study unit, ARI3129, will explore these topics in detail. By creating new artificial examples that capture the variability and complexity of real-world data, data augmentation can help make the models more robust and effective in various computer vision applications.

Overall, data augmentation with photometric and geometric distortions is a powerful tool in machine learning. It has been widely used in computer vision research and practice to improve the performance of deep learning models.

Part A

This part of the project introduces students to conducting background research on a topic. The first question focuses on writing background research on data augmentation in computer vision. In this part of the assignment, students are encouraged to refrain from using GenAI since it will be used in Question 2. Using it for this question might introduce unnecessary confusion and complexity in the other question while impacting what you can write in that question.

The second encourages students to use GenAI to assess the quality of their writing while helping them reflect on their work.

Q1: Writing a background about Data Augmentation

Each team member is to write a 400-word (minimum) background about data augmentation in computer vision. Include an overview of computer vision milestones, such as developing convolutional neural networks and large-scale image datasets like ImageNet. Discuss the relevance of data augmentation in deep learning models for computer vision and describe the different data augmentation techniques used, such as flipping, rotation, scaling, cropping, and colour jittering. Additionally, provide a detailed discussion of recent advancements in data augmentation techniques, such as Cutout, Mixup, and AutoAugment, and explain how these techniques can help improve the performance of deep learning models. Finally, describe how data augmentation has been used in specific computer vision applications, such as object detection and image segmentation, and provide examples of recent papers using data augmentation to achieve state-of-the-art results, such as the bag of freebies in the YOLOv4 paper. Proper citations and references are required throughout the background section.

Deliverable: Each student must submit a 400-word background, including references, in the documentation without the GenAI refinement.

Maximum of 30 Marks, organised accordingly:

- Introduction (5 marks): Provide a clear and concise introduction to the topic, highlighting the importance of data augmentation in computer vision.
- Historical Context (5 marks): Overview of computer vision milestones.
- Data Augmentation Techniques (10 marks): Description of different data augmentation techniques used in deep learning models for computer vision, including flipping, rotation, scaling, cropping, and colour jittering.
- Recent Advancements (5 marks): Discuss recent advancements in data augmentation techniques.
- Application in Computer Vision (10 marks): Provide a description of how data augmentation has been used in specific computer vision applications. At least 12 references should support this background.
- Citation and References (5 marks): Proper citations and references throughout the background section, including a reference list at the end.

Q2: Using GenAI as a writing assistant

For this question, students are encouraged to use GenAI as a writing assistant to help them improve the quality and effectiveness of their writing and generate new ideas and approaches to the project. The following guidelines apply to this question:

- Copy and paste the text of the background section into the chat window with GenAI.
- Ask GenAI for suggestions on improving the writing, such as grammar and punctuation corrections, rephrasing sentences or paragraphs, and suggestions for better word choices.
- Ask GenAI for feedback on the overall structure and flow of the writing, as well as suggestions for any additional content that could be added to improve the quality of the writing.
- Ask GenAI to act as a strict peer-reviewer and give you the pros and cons of your work.

Deliverables:

1. Provide a set of prompts (minimum of 5) used and the replies given by GenAI to demonstrate how you got feedback about your writing.
2. Each Student is to submit a brief 100-word reflection of what was learnt from the responses of GenAI.

Maximum of 10 Marks, organised accordingly:

- 5 marks for using and reporting prompts/responses.
- 5 marks for the set of individual reflections

End of Part A

Part B

This part of the project focuses on the practical aspect of data augmentation. Students will be required to investigate how data augmentation is available through the mainstream frameworks while also implementing a selection of techniques from first principles

For this project, students will work with the COTS dataset[1], which is freely available through the website (www.cotsdataset.info). The COTS dataset contains scenes with an incremental nature, and it is commonly used in computer vision research for object detection and tracking tasks.

Q3: Data Augmentation in mainstream machine learning frameworks

Data augmentation in TensorFlow and PyTorch is achieved through specialised libraries and functions that can be used to manipulate image data and generate new augmented images. In this question, you are required to document how data augmentation is achieved in both TensorFlow and PyTorch. You must demonstrate how this is achieved in both frameworks by creating a Jupyter Notebook for each framework and generating a number of samples on the COTS dataset.

Using TensorFlow and PyTorch, generate data augmentation on the COTS dataset. Implement at least three image transformations using the ImageDataGenerator class in TensorFlow and the transforms module in PyTorch. Generate at least 5 augmented images per input image. Save the augmented images to a new directory on your local machine. Ensure that the saved images are in a compatible format (such as JPEG or PNG) and that they are named in an easy way to identify. Make use of comparative metrics of your choice to demonstrate how different each of the resultant images is from the original image.

Include a brief write-up that explains the steps you took to generate the data augmentation and save the output to file, as well as any challenges you encountered during the process.

Deliverables:

[1] D. Seychell, C. J. Debono, M. Bugeja, J. Borg and M. Sacco, "COTS: A Multipurpose RGB-D Dataset for Saliency and Image Manipulation Applications," in IEEE Access, vol. 9, pp. 21481-21497, 2021. Paper available through: https://ieeexplore.ieee.org/document/9340352

1. A Jupyter Notebook demonstrating data augmentation in TensorFlow
2. A Jupyter Notebook demonstrating how data augmentation is done in PyTorch
3. A zip file containing 2 folders for the output of both frameworks demonstrated on a random 5 images of the COTS dataset.
4. Documentation showing the team's process to achieve the above results while declaring any use of generative AI.

Maximum of 20 Marks, organised accordingly:

- Implementation of Data Augmentation in TensorFlow and PyTorch (15 marks):
    - Use of ImageDataGenerator class and transforms module to implement at least three image transformations
    - Generation of at least 5 augmented images per input image
    - Proper handling of image data, including loading, processing, and saving to file
    - Efficient use of memory and computational resources
- Quality of Augmented Images (5 marks):
    - Images are visibly different from the original images, indicating successful augmentation
    - Augmented images are in a compatible format and named in a way that is easy to identify

Q4: Implementation of Data Augmentation Techniques

Data augmentation can be achieved through various transformations, including photometric distortions and geometric distortions. Consider the following distortions:

Photometric Distortions:
- Brightness variation
- Contrast variation
- Color channel shifts
- Saturation variation
- Gaussian noise

Geometric Distortions:
- Rotation
- Scaling
- Translation
- Shearing
- Flipping

Using OpenCV, implement a program in a Jupyter Notebook that applies photometric and geometric distortions to the same subset of images used in Question 3. Apply at least three

different techniques and three geometric distortions for photometric distortions. Implement these techniques from first principles, hence not calling a function that simply carries out the distortion.

For each distortion, generate 5 augmented images per input image. Compare the results of the output of your implementation with similar techniques

Deliverables:

1. A Jupyter Notebook containing the implementation of the 6 distortion techniques
2. A zip file containing the output of each technique on a random 5 images of the COTS dataset.

Maximum of 20 marks, organised accordingly:

- Implementation of Photometric Distortions in OpenCV (6 marks):
  - Implementation of at least three photometric distortion techniques
  - Generation of at least 5 augmented images per input image
- Implementation of Geometric Distortions in OpenCV (6 marks):
  - Implementation of at least three geometric distortion techniques
  - Generation of at least 5 augmented images per input image
- Quality of Augmented Images (4 marks):
  - Images are visibly different from the original images, indicating successful augmentation
  - Augmented images are in a compatible format and named in a way that is easy to identify
- Comparison of your implementation with a framework (4 marks)
  - Reporting of similarity/difference using a set of metrics of your choice

---

End of Part B

Generative AI Journal Guidelines

The objective of this document is to critically examine and reflect on how generative AI models, like ChatGPT or Gemini, were used in this project. This documentation will focus on ethical considerations, the methodology employed, and the specific contributions of generative AI to improving your work. You must cite any references or resources you consulted about generative AI.

The document, which is separate from the technical paper, has a page limit of 10 pages, including references and should contain the following sections:

1. **Introduction**

   a. Briefly describe the generative AI models you used and the rationale behind that choice. (Maximum of 1 page)

2. **Ethical Considerations**
   a. Discuss the ethical aspects of using generative AI in your project. This should include issues like data bias and privacy. (Maximum of 1 page)

3. **Methodology**
   a. Outline the methods and steps to integrate the generative AI model into your work.

4. **Prompts and Responses**
   a. List the specific prompts that were used with the generative AI model. For each prompt, also include the generated response and explain how it contributed to improving your project.

5. **Improvements and Contributions**
   a. Discuss the specific areas where generative AI enhanced your work. This can include, but is not limited to, data analysis, formulation of ethical considerations, literature review enhancement, or idea generation.

6. **Individual Reflection**
   a. Reflect on your personal experience using generative AI in your project. Discuss what you learned, what surprised you, and how your perspective on using AI in academic projects has changed, if at all.

7. **References and List of Resources Used**

End of Project Specifications