DMP title

Project Name Deep Learning Methods for Global Illumination Transport in Computer Graphics - DMP title

Grant Title C14/21/073

Principal Investigator / Researcher Philip DutrÃ©

Description The project develops algorithms using neural networks to learn how to generate synthetic images. Data used in the project: - developed computer algorithms - input training data (images), generated by computer programs - output data (images), generated by computer programs

Institution KU Leuven

1. General Information

Name of the project lead (PI)

Philip Dutré

Internal Funds Project number & title

Learning Methods for Global Illumination Transport in Computer Graphics
Project reference nr:C14/21/073
SAP project code: 3E210528
Supervisor: Philip Dutré
Fellow:
Funder: BOF
Programme: C1-project 4 Jaar
Project start date: 1/10/2021
Project end date: 30/09/2025
C1
C14/2

2. Data description

2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

1. Images for training data set for neural networks. This is part of the actual research work in an interative cycle. All images are generated by computer programs (global illumination renderers), so images are generated as needed on-the-fly. Estimated size per image (.exr formats): 512x512 resolution, 3MB per image, order of magnitude 10K images ==> 30GB
2. Program code, version controls, generated neural networks ==> 1 GB
3. Output images, variable amount, same data formats as in 1 ==> 30 GB

3. Ethical and legal issues

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

NA

3.2. Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s).

NA

3.3. Does your research possibly result in research data with potential for tech

transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

NA at time of writing. This could be a possibility, but not factored in at this point.

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

No.

4. Documentation and metadata

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

All data will consist of computer-geenrated images or computer program code.

The images are generated using implemented algorithms, and therefore, the desciription of how the images are generated is contained in the computer code. The amount, time, resolution, ... of the images is recorded withe each speciic image.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

Metadata for computer generated images is incorporated in the file formats themselves (date, resolutuion, size, ...)

5. Data storage and backup during the project

5.1. Where will the data be stored?

Various platforms:

1. Local computers of researcher during research work
2. Cloud-based platforms for storage, linked to evrsion control systems for software (e.g. Github)
3. Back-up systems of Department of Computer Science.

5.2. How will the data be backed up?

The Department of Computer Science provides backup functionality, which will be used.

5.3. Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

Yes, the Department of Computer Science has extensive backup functionality. Given the limited amount of data, we do not foresee significant problems.

2nd tier are cloud-based services.

5.4. What are the expected costs for data storage and backup during the project? How will these costs be covered?

Costs are managed by the Department of Computer Science. Internal cost systems are used to have every research group contribute from project overheads etc. Given the limited amount of data, we do not foresee any problems.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Internal storage Dpeartment of Computer Science.

Data is not critical or provacy-sensitive, so this is less of an issue.
In case of breach, data can always be regenerated if necessary.

6. Data preservation after the end of the project

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained. This is the current practice in my fioeld (computer graphics).
Data is usually included as part of project websites or supplemental materials for published papers as well.

6.2. Where will these data be archived (= stored for the long term)?

1. University (Department of Computer Science) servers
2. Cloud-based servers.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

University (Department of Computer Science) servers: costs are carried by the department - each research group contributes a sum per capita per year.

7. Data sharing and re-use

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

No. All data is generated by the project itself.

7.2. Which data will be made available after the end of the project?

All generated images and computer code. This is common practice in our field.

7.3. Where/how will the data be made available for reuse?

- In an Open Access repository

7.4. When will the data be made available?

- Immediately after the end of the project

Data might be released during the project, as part of published papers or technical reports. This is common practice in the research discipline.

7.5. Who will be able to access the data and under what conditions?

Data will be made open access, downloadable by all interested parties.

7.6. What are the expected costs for data sharing? How will these costs be covered?

NA - through website and cloud-based servers.

8. Responsibilities

8.1. Who will be responsible for the data documentation & metadata?

Researcher employed by the project + PI Philip Dutré

8.2. Who will be responsible for data storage & back up during the project?
Researcher employed by the project + PI Philip Dutré

8.3. Who will be responsible for ensuring data preservation and sharing?
Researcher employed by the project + PI Philip Dutré

8.4. Who bears the end responsibility for updating & implementing this DMP?
The end responsibility for updating and implementing the DMP is with the supervisor (promotor).