

# Creative AI: artificial intelligence techniques for image/audio/video creation and generation

Amal El Fallah Seghrouchni, Omar Elharrouss

---

**Key-words:** Image creation, text-to-image generation, sketch-to-image generation, panoramic image creation.

## 1 Research Motivation

The creation of an image from another and from different types of data including text, scene graph, and object layout, is one of the very challenging tasks in computer vision. In addition, capturing images from different views for generating an object or a product can be exhaustive and expansive to do manually. Now, using deep learning and artificial intelligence techniques, the generation of new images from different type of data is became possible. For that, a significant effort has been devoted recently to develop image generation, construction and reconstruction strategies with a great achievement. With the introduction of artificial intelligence and deep learning techniques, working on images and videos becomes easier for recognition, segmentation, and prediction tasks[1]. In addition to this, these models give the researcher the ability to create new tasks like image transformation, image/video generation, and new image/video construction. This helped to create art as well as playing games with high performance. These tasks are in general new, which makes the development of novel techniques a trend [2]. For that, the creation of new images/video for different purposes, a challenging and attractive subject.

## 2 Problem Statement

Generating new image/video is one of challenging task in computer vision. With the introduction of deep learning techniques especially using CNN models this task become doable which attracted the researcher to create new images not just based on other images but also from text, sketch, scene graph, layout or also from a set of scenes [3]. Few years ago, the development of AI techniques including generative adversarial networks (GANs) has revive images generation task to generate new images with a high performance in term of image quality and pertinent content[4]. Its have produced realistic images of human face, furniture, or scene that are difficult to distinguish from real images [5]. Consequently, AI models have been used to create target images using different ways. Many works have been proposed exploiting various features and techniques to reach the best and effective results.

### 3 Research Scope

The aim of this thesis is to propose and develop deep-learning-based methods from different types of data. For that, a literature review should be done to understand different directions and existing approaches. Then, proposing new deep-learning-based methods for Image/video creation based on the nature of the adopted algorithms, type of data used, and main objective.

### 4 Admission Criteria

The PhD position is proposed by the International Center of Artificial Intelligence of Morocco, of the Mohammed VI Polytechnic University. Applicants with excellent curricula must be holders of a Master's, an engineering or an equivalent recognized degree in Computer Science or Applied Mathematics. In addition, they should have skills in Programming (Python and C++) and good communication skills in English. Particular attention will be given to the suitability of this research project with the applicant's background.

### References

- [1] Tingting Qiao, Jing Zhang, Duanqing Xu, and Dacheng Tao. Learn, imagine and create: Text-to-image generation from prior knowledge. *Advances in Neural Information Processing Systems*, 32:887–897, 2019.
- [2] Janet Rafner, Arthur Hjorth, Sebastian Risi, Lotte Philipsen, Charles Dumas, Michael Mose Biskjær, Lior Noy, Kristian Tylén, Carsten Bergenholtz, Jesse Lynch, et al. crea. blender: A neural network-based image generation game to assess creativity. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, pages 340–344, 2020.
- [3] Jiuxiang Gu, Handong Zhao, Zhe Lin, Sheng Li, Jianfei Cai, and Mingyang Ling. Scene graph generation with external knowledge and image reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1969–1978, 2019.
- [4] Bowen Li, Xiaojuan Qi, Thomas Lukasiewicz, and Philip HS Torr. Controllable text-to-image generation. *arXiv preprint arXiv:1909.07083*, 2019.
- [5] Naser Damer, Fadi Boutros, Khawla Mallat, Florian Kirchbuchner, Jean-Luc Dugelay, and Arjan Kuijper. Cascaded generation of high-quality color visible face images from thermal captures. *arXiv preprint arXiv:1910.09524*, 2019.