AI-Assisted Creative Expression: a Case for Automatic Lineart Colorization

Yliess Hati

January 26, 2023

Contents

Abstract . . . 1
Introduction . . . 2
    Motivations . . . 2
    Problem Statement . . . 2
    Contributions . . . 3
    Concerns . . . 3
    Outline . . . 3
Background . . . 4
    History of Artificial Intelligence . . . 4
    Neural Networks . . . 4
    Autoencoders . . . 4
    Variational Autoencoders . . . 4
    Generative Adversarial Networks . . . 4
    Denoising Diffusion Models . . . 4
Contrib I (Find Catchy Explicit Name) . . . 5
    State of the Art . . . 5
    Method . . . 5
    Setup . . . 5
    Results . . . 5
    Summary . . . 5
Contrib II (Find Catchy Explicit Name) . . . 6
    State of the Art . . . 6
    Method . . . 6
    Setup . . . 6
    Results . . . 6
    Summary . . . 6
Contrib III (Find Catchy Explicit Name) . . . 7
    State of the Art . . . 7
    Method . . . 7
    Setup . . . 7
    Results . . . 7
    Summary . . . 7
Contrib IV (Find Catchy Explicit Name) . . . 8
    State of the Art . . . 8
    Method . . . 8
    Setup . . . 8
    Results . . . 8
    Summary . . . 8

Ethical and Societal Impact 9
Conclusion 10
References 11

Abstract

Introduction

Humans possess the ability to perceive and understand the world allowing us to accomplish a wide range of complex tasks through the combination of visual recognition, scene understanding, and communication. The ability to quickly and accurately extract information from a single image is a testament to the complexity and sophistication of the human brain and is often taken for granted. One of the Artificial Intelligence (AI) field's ultimate goals is to empower computers with such human-like abilities, one of them being creativity, being able to produce something original and worthwhile [6].

Computational creativity is the field at the intersection of AI, cognitive psychology, philosophy, and art, which aims at understanding, simulating, replicating, or in some cases enhancing human creativity. One definition of computational creativity [7] is the ability to produce something that is novel and useful, demands that we reject common beliefs, results from intense motivation and persistence, or comes from clarifying a vague problem. Top-down approaches to this definition use a mix of explicit formulations of recipes and randomness such as procedural generation. On the opposite, bottom-up approaches use Artificial Neural Networks (ANNs) to learn patterns and heuristics from large datasets to enable non-linear generation.

We, as a species, are currently witnessing the beginning of a new era where the gap between machines and humans is starting to blur. Current breakthroughs in the field of AI, more specifically in Deep Learning (DL), are giving computers the ability to perceive and understand our world, but also to interact with our environment using natural interactions such as speech and natural language. ANNs, once mocked by the AI community [5], are now trainable using Gradient Descent (GD) [8] thanks to the massive availability of data and the processing power of modern hardware accelerators such as Graphical Processing Units (GPUs), Tensor Processing Units (TPUs), and Neural Processing Units (NPUs).

Neural Networks (NNs), those trainable general function approximators, gave rise to the field of generative NN. Specialized DL architectures such as Variational Autoencoders (VAEs) [4], Generative Adversarial Networks (GANs) [2], Denoising Diffusion Models (DDMs) [3], and Large Language Models (LLMs) [1, 9] are used to generate artifacts such as text, audio, images, and videos of unprecedented quality and complexity.

This thesis aims at exploring how one could train and use generative NN to create AI-powered tools capable of enhancing human creative expression.

Motivations

- A case for Lineart Colorization

Problem Statement

- Black & White Lineart VS Gray Scale
- Incomplete Information Challenge fo Computer Vision
- Natural Artisitic Control Back to the User

Contributions

- Recipe for curating datasets for the task of automatic colorization
- 3 Models exploring different aspect of the topic:
  - PaintsTorch: High Quality, User-Guided, Fast Realtime Feedback
  - StencilTorch: Human-Machine Collaboration, Human-in-the-Loop
  - StableTorch: Variance and Iterative Exploration
- A reflexion on Current Generative AI Ethical and Societal Impact in our Society

Concerns

- Raise awareness about
  - Deepfakes
  - Model Fabulations
  - Ownership & Copyright Ambiguities
  - Biases & Discrimination
- About this work
  - Images used only for Educational and Research Purposes
  - Only describe recipes for reproducibility
  - Dataset and Weights are not Distributed (Only Code)

Outline

- Plain Language Expanded TOC

Background

History of Artificial Intelligence

Neural Networks

Autoencoders

Variational Autoencoders

Generative Adversarial Networks

Denoising Diffusion Models

Contrib I (Find Catchy Explicit Name)

State of the Art

Method

Setup

Results

Summary

Contrib II (Find Catchy Explicit Name)

State of the Art

Method

Setup

Results

Summary

Contrib III (Find Catchy Explicit Name)

State of the Art

Method

Setup

Results

Summary

Contrib IV (Find Catchy Explicit Name)

State of the Art

Method

Setup

Results

Summary

Ethical and Societal Impact

Conclusion

References

[1] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*. 33, (2020), 1877–1901.

[2] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. 2014. Generative adversarial nets[1]. *Advances in neural information processing systems* (2014).

[3] Ho, J., Jain, A. and Abbeel, P. 2020. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*. 33, (2020), 6840–6851.

[4] Kingma, D.P. and Welling, M. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*. (2013).

[5] Le Cun, Y. 2019. *Quand la machine apprend: La révolution des neurones artificiels et de l'apprentissage profond*[2]. Odile Jacob.

[6] Mumford, M., Medeiros, K. and Partlow, P. 2012. Creative thinking: Processes, strategies, and knowledge. *The Journal of Creative Behavior*. 46, (Mar. 2012). DOI:https://doi.org/10.1002/jocb.003[3].

[7] Newell, A., Shaw, J.C. and Simon, H.A. 1959. *The processes of creative thinking*[4]. RAND Corporation.

[8] Rumelhart, D.E., Hinton, G.E. and Williams, R.J. 1986. Learning representations by back-propagating errors. *Nature*. 323, 6088 (Oct. 1986), 533–536. DOI:https://doi.org/10.1038/323533a0[5].

[9] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*. 30, (2017).