Applied Deep Learning [194.077]

Image Colorization of Archived Images

Anastasia Cissa, ID: 11937948

TECHNISCHE
UNIVERSITÄT
WIEN

Course: Applied Deep Learning [194.077]
Academic Year: WS2024
January 25, 2025

Contents

1 What is the problem that you tried to solve?

The problem tackled in this project is the automation of image colorization for historical black-and-white photographs. Specifically, the goal was to apply deep learning models to transform grayscale photos into realistic colorized images. This project aimed to preserve and enhance the cultural heritage of Moldova by restoring archival photographs and making them more visually engaging. The process intended to reduce the manual effort and time required for such restoration tasks.

2 Why is it a problem?

Manual colorization of historical photographs is a time-consuming, expensive, and labor-intensive process. It requires artistic skills, a deep understanding of historical contexts, and considerable patience. In Moldova, no automated tools have been previously utilized for this purpose. As a result, all the people interested in the process of historical pictures restoration don't have a tool that can be applied for such problem. This limitation makes it challenging to share and celebrate the country's cultural heritage by a broader audience. By automating the process, this project aims to bridge the gap, enabling faster restoration, enhancing accessibility, and promoting cultural preservation on a larger scale.

3 What is your solution?

The solution involves developing and applying deep learning models to automate the process of image colorization. The core methodology leverages a U-Net generator combined with a PatchGAN discriminator within a Conditional Generative Adversarial Network (cGAN) framework. This model was designed to:

- Accept grayscale images (L channel in L*a*b* color space) as input.
- Predict and generate the missing color channels (a*b* channels).
- Produce realistic, high-quality colorized images that align with the historical context of the original photographs.

The approach for this project is based on the paper *Image-to-Image Translation with Conditional Adversarial Networks* by Isola et al., which introduced the pix2pix framework. This framework uses a Conditional Generative Adversarial Network (cGAN) to map input images to corresponding output images. The U-Net generator in this architecture includes skip connections, enabling the fusion of low-level and high-level features for better structural preservation. Meanwhile, the PatchGAN discriminator ensures local realism by evaluating small patches of the generated images, making it particularly effective for tasks like colorization, where local details matter.

Additionally, the project was inspired by the implementation provided in `mberkay0/image-colorization` on GitHub. This implementation provided practical insights into adapting the pix2pix framework for image colorization tasks. While the GitHub repository primarily served as a starting point, the methodology was further customized for the specific requirements of this project.

By adapting these frameworks, the project combines the strengths of the U-Net generator and the PatchGAN discriminator to address the specific challenges of image colorization for Moldovan archival photos. The principles of balancing adversarial loss for realism and reconstruction loss for accuracy were integral to the model's design and training process.

The project utilized a blend of pre-existing datasets like Places and ImageNet for training and a custom dataset of Moldovan archival photos for fine-tuning and evaluation. This hybrid approach ensured that the model was both robust and domain-specific.

4 Why is it a solution?

Deep learning offers a scalable and efficient solution for the image colorization problem because it can learn complex mappings from grayscale to color directly from data. Key advantages of this approach include:

- **Automation**: The trained model can process hundreds of images in a fraction of the time required for manual colorization.
- **Consistency**: The model produces uniform and reproducible results, eliminating human bias and variability.
- **Scalability**: Once trained, the model can be deployed to process entire collections of archival photos without additional effort.

However, the project also revealed some limitations. While the model performs well on general photos, it struggles with out-of-distribution data, such as rare historical contexts or unique cultural elements specific to Moldova. Additionally, unrealistic colorization in some scenarios highlighted the need for further fine-tuning and more diverse training data. Despite these challenges, the approach represents a significant improvement over manual methods.

5 Main Takeaways and Insights

5.1 Key Insights:

1. **Model Architecture**: The U-Net generator, with its skip connections, significantly improved the preservation of structural details in the grayscale input. The PatchGAN discriminator ensured that the generated images were locally realistic, enhancing the overall quality of colorization.

2. **Training Strategy**: Pretraining the generator with an L1 loss before adversarial fine-tuning was crucial for stabilizing training. This two-stage approach reduced training time and improved model performance.

3. **Dataset Limitations**: The lack of diversity in the training data limited the model's ability to generalize to unique historical contexts. Photos representing Moldovan traditions, customs, and specific cultural elements posed significant challenges for the model.

4. **Unexpected Behaviors**: The model occasionally applied excessive magenta and green tones in unexpected regions, revealing potential biases in the training data. These artifacts emphasized the importance of dataset curation and balancing.

5.2 Lessons Learned:

- **Data Preprocessing**: Cleaning and preparing the dataset was more time-intensive than anticipated, especially ensuring that the images were of consistent quality and relevance.

- **Evaluation Challenges**: Assessing the model's performance required both quantitative metrics (e.g., L1 loss) and qualitative analysis (visual inspections). Striking the right balance between these evaluation methods was critical.

- **Tool Selection**: While the PastVu API and Instagram scraping were effective for data collection, manual annotation and verification required significant effort.

- **Deployment**: Preparation of the model can become challenging if model's architecture includes several steps and is saved during each step of its training. It becomes difficult to incorporate each step of the training.

6 If You Would Do the Same Project Again

If I were to redo this project, I would:

- **Expand the Dataset**: Include a broader and more diverse range of historical photos, particularly those representing Moldovan traditions and landmarks, to improve the model's generalization. In order to do that, I would contact many institutions long before, so we would have time to address all potential problems of the project and our collaboration.
- **Improve Evaluation Metrics**: Develop custom evaluation metrics tailored to the model except loss functions in order to have more performance metrics.
- **Automate Model Preparation**: Enhance the code so models could be deployed easily in different ways.

7 Time Spent on the Project

7.1 Reasons for Underestimation:

- **Dataset Collection**: The manual effort required to clean, annotate, and verify the dataset was higher than expected, especially when ensuring its relevance to Moldovan culture.
- **Model Training**: Debugging and addressing issues like overfitting and color imbalance extended the training timeline.
- **Documentation and Presentation**: Preparing clear and comprehensive documentation, along with visualizations, took additional time to ensure quality.
- **Deployment**: Preparing application that would work effortlessly required quite al ong time due to long time preparation of the model that works in 2 steps. It became a challenge to incorporate both steps of the model.

8 Conclusion

This project demonstrated the potential of deep learning for automated image colorization, providing an efficient and scalable solution for preserving historical photographs. While the results were promising, the project also highlighted the importance of domain-specific data. Future iterations should focus on enhancing dataset diversity, improving evaluation metrics. Overall, this work represents a meaningful step toward modernizing the preservation of Moldova's cultural heritage.