Practical machine learning and deep learning course project proposal

Automated manga cleaning using CNNs

Team:
Vyacheslav Sergeev
vy.sergeev@innopolis.university
Nikita Strygin
n.strygin@innopolis.university
Dinislam Gabitov
d.gabitov@innopolis.university

Introduction & Problem context

Nowadays japanese manga has become a very prominent form of entertainment with up to 70% of comics being sold in some European countries being manga. However with this popularity came the challenge of accessibility. Since overwhelming majority of manga is produced in Japan the European readers have to seek the translated versions. Since the number of manga importers is still pretty limited unofficial translations of manga have become commonplace and are used by a large portion of manga readers.

The challenge of translating the manga can easily be underestimated. The most important thing to keep in mind is that not only the translator has to translate the japanese text to english, which is somewhat easy, but also to insert the text back Into the manga. This presents an unusual challenge since unofficial translators of manga have to first scan the manga and then remove all of the text in the scanned images by hand, then reinsert the text back into the image and make it look like a part of the original manga.

Let us summarize the process for translating the manga in a few general steps:
1. Scan the manga and enhance it
2. Remove all of the text
3. Translate the text and reinsert it back
4. Position the text in a way to fit the setting
5. Choose the font for the text to make it look good

Here is an example of what the process of translating manga looks like in a few steps:

Aims & Goals of the project

We have decided to implement a solution for at least the first two steps of manga translation colloquially known as manga cleaning. Since the tool that is most commonly used for cleaning is adobe photoshop we hope to develop a plugin for it that would allow automagic cleaning in a press of a button.

Goals can be summarized as follows:
1. Create a solution that would allow cleaning of different manga in a more or less stable way
2. (Optional) Wrap the solution into a photoshop plugin
3. (Highly optional) Develop the capability to fill in the space left from removing text that looks more or less natural
4. (Extremely unlikely) Add a capability to select/generate the proper font for the translated text

An example of what we will strive to achieve:

The ideal solution would also be able to deal with difficult cases such as this one:
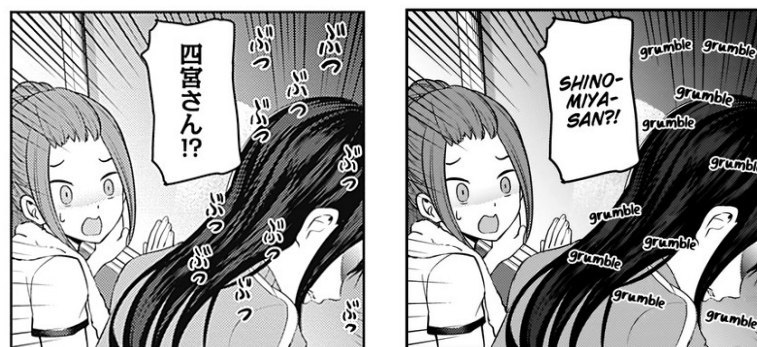
In this case, the program has to identify drawn sound effects which look notably different from the text in speech bubbles and then after deleting them redraw the segments which were obsured.

Proposed architecture

Most of the research suggests the usage of different flavours of U-Net along with different kinds of pre- and post- processing.

Data availability

The biggest challenge for this project will be data availability. We have already begun contacting different communities who might have some data that we could use. There is also manga109 dataset which could be very useful (http://www.manga109.org/en/) however its uses are pretty constrained and the tool is restricted in access. If the pre-made datasets will not be available there are hopes of data augmentation and manual data generation. It is certainly not preferred, but could be done.

Competition & existing solutions

As far as we know there are no existing solutions that would work well enough to gain any use. There is a tool located at this url:http://codensuch.com/speech-bubble-ai/ai-manga-translation-part3 however it still suffers from issues including under-identifying speech bubbles.
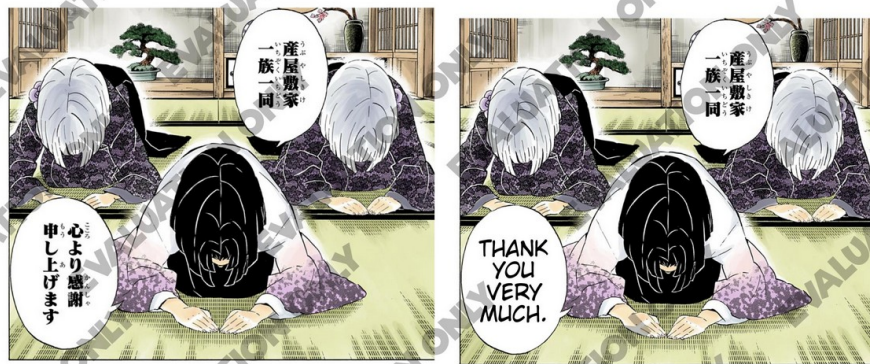
Example of under-identifying a speech bubble

Possible challenges

Possible challenges include but are not limited to:
- Lack of data
    - Hard to find clean manga in general
    - Challenging to find manga that was cleaned of SFX consistently
    - Hard to find consistent datasets in general
- Bad data quality
    - Undererased text in the data
    - Manga bent while scanning
    - Bad image quality
    - Images could be augmented by filters
- Hard to make model generalise to different styles of manga
- Generalising to both text and SFX is hard
- Learning the model may be too computationally intensive

Papers & references

Ko, U-Ram & Cho, Hwan-Gue. (2020). SickZil-Machine: A Deep Learning Based Script Text Isolation System for Comics Translation. 10.1007/978-3-030-57058-3_29.
Available at: https://www.researchgate.net/publication/343646364_SickZil-Machine_A_Deep_Learning_Based_Script_Text_Isolation_System_for_Comics_Translation

Juli ́ an Del Gobbo. (2020). Unconstrained Text Detection in Manga. Available at: http://dc.sigedep.exactas.uba.ar/media/academic/grade/thesis/JDelGobbo_tesis.pdf