In the project, the authors consider the task of generating images in auto-regression ways, namely line (sequence) by line. The obvious restriction of such an approach lies in the length of the considered sequence, as modern recurrent architectures like RNN and GPT fail working with long sequences. The project's author proposed applying different tokenization techniques, similar to using it for NLP tasks.

The project statement and related work are pretty well described in the report. The main idea is understandable.
More particular comments are the following:

1. In the Introduction, the connection between lines 45 --- 47 is not obvious to me. I can guess that your approach uses ideas from NLP tasks applied to synthetically generated sequences. However, it will be better if you provide a more gentle connection. In addition, you should add some very common details about considered tokenization methods. Are they well-known (so you should cite some work) or proposed by you (still need some citation about how you came up with this solution)? At least BPE is well-known, so you should already cite the corresponded paper in the Introduction.
2. The Section Experiments suffers from underfilling now. In section 3.1, it will be more beneficial to provide a reasonable conclusion about Figure 2. At what point is enough? Is it good or bad? In addition, please, "sign" the axes in this Figure.
3. A similar comment is about Section 3.2. You should provide some analysis of the obtained picture. Of course, some of the "insights" are obvious; nevertheless, you should describe them.
4. You should provide the link to GitHub in the text of the report.

Minor comments:

1. It's nice to put the report's overall structure at the end of the Introduction.
2. A caption should be more informative. E.g., what is the source of Figure 1? Does it present one of the worst examples generated by your model or by somebody else?
3. Lines 87, please, provide the link/citation on the CELEBA dataset.
4. Lines 89--90, please, put "\n" on one line.
5. Line 97, the reference is displayed incorrectly.
6. It will be better to proofread the text one more time. There are some grammar mistakes.

As for the practical part, the experiment protocol seems reasonable. I also suggest using a quantitative approach to evaluate the method (now, it's only qualitative), e.g. Friche Inception Distance, Inception Score (for GAN evaluation). Please, provide the necessary requirements for running the notebook.