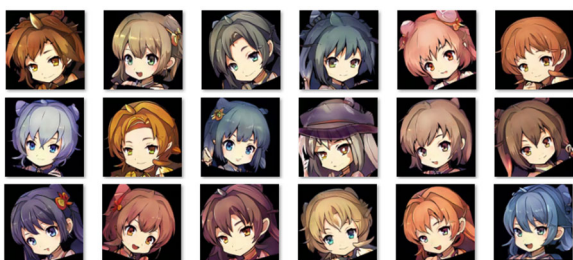


StyleGAN2 に基づくゲームキャラクターの顔画像生成における生成画像の品質制御方式 A Quality Control Method of Generated Image of Facial Image Generation of Game Character based on StyleGAN2

周 逸峰[†] 阿倍 博信[†]
Yifeng Zhou Hironobu Abe

1. はじめに

キャラクターを動かすゲームの開発においては、大量の顔画像が必要になる。近年、AI を用いて顔画像を自動生成するサービスが増加しているが、我々の開発チームにおいても、ゲームへの素材の提供を想定して、StyleGAN2[1]に基づく表情差分付きのゲームキャラクターの顔画像生成システムの開発に取り組んでいる[2]。図 1 に本システムにより生成された顔画像の一例について示す。



©2014 CloverLab.,Inc.

図 1 顔画像生成システムにおける生成画像の例

図 1 において、システムが生成した画像の中で、一定数の異常画像が存在することが課題となっていた。そこで、本研究では、機械学習を用いて生成した顔画像を正常画像と異常画像を分類することにより、StyleGAN2 で生成したゲームキャラクターの顔画像を品質制御する方式について提案する。また、提案方式について評価実験を行い、その有効性について確認する。

2. 関連研究

2.1 マルチクラス画像分類

マルチクラス画像分類とは、クラスが複数存在し、サンプルに対してどれかのクラス 1 つに対して割り当てるように最適化する問題である。MNIST や CIFAR などはこれにあたる。

CNN に基づく画像分類方式の実装として、TensorFlow が提供する Image classification[2]がある。本方式が提供する Sequential モデルでは、花の画像分類を行うチュートリアルを提供している。Sequential モデルは、CNN をはじめとする層の順序が明確なモデルに適しており、また、過学習を防ぐための具体的な手法、データ拡張やドロップアウトの技術についても提供されている。

2.2 マルチラベル画像分類

マルチラベル画像分類とは、通常の画像ごとに 1 つのラベルのみが割り当てられるマルチクラス画像分類とは異なり、与えられた画像からラベルのセットを予測する方式である[3]。1 つの画像に対して複数のラベルがある方が自然であり、このような方式はより幅広い用途に適用できる。

[†] 東京電機大学 Tokyo Denki University

3. 生成画像の品質制御方式の提案

前述の通り、本研究では、機械学習を用いて生成した顔画像を正常画像と異常画像を分類することにより、StyleGAN2 で生成したゲームキャラクターの顔画像を品質制御する方式について提案する。正常画像と異常画像の分類方式としては、関連研究を参考にシンプルなマルチクラス画像分類とマルチラベル画像分類の 2 方式について検討していく方針とする。

3.1 マルチクラス画像分類

マルチクラス画像分類に基づく正常画像と異常画像の分類を目的として、TensorFlow が提供する Sequential モデルを用いた事前評価を実施した。

まず、参考文献[2]に示した顔画像生成システムが生成した顔画像 1,000 枚を対象として、手作業によりアノテーション作業を行った。その結果、図 2 に示した通り、10%の画像は異常画像として分類されるという結果となった。



©2014 CloverLab.,Inc.

図 2 異常画像の例

次に、TensorFlow が提供する Sequential モデルを用いてモデル構築を行い、有効性の評価を行った結果、異常画像の検出率は 98%となった。しかし、実際にアノテーションを行う際には、正常、異常の判断基準が難しい、などの問題があるため、方式の改良を検討することにした。

3.2 マルチラベル画像分類

次に、3.1 の実施結果を踏まえ、主に正常、異常の判断基準の明確化を目的としてマルチラベル画像分類に取り組むこととした。

3.2.1 異常特徴の抽出による判断基準の設定

まず、顔画像自動生成システムが生成した画像に対する分析を行い、正常、異常の分類と、異常については異常特徴の分類を行った。その結果、異常特徴として、①白い目(eye_ball)、②瞳孔拡散(eye_pup)、③髪に細長いもの付き(hair_bow)、④顔両側に細長いもの付き(hair_side)、⑤前髪崩れ(hair_fro)、⑥後ろ髪不自然(hair_back)、という 6 つの異常特徴を抽出した。今回抽出した異常特徴の例について図 3 に示す。

また、正常、異常の判断基準の個人差を吸収することを目的として、抽出した異常画像を提示する形で 63 名(大学生:15 名、オープンキャンパス来場者:48 名)を対象にアンケート調査を行った。その結果を表 1 に示す。



©2014 CloverLab.,Inc.

図 3 抽出した異常画像の例

表 1 アンケート結果

異常の特徴	判断回数	割合 (%)
① eye_ball	38	60.3
② eye_pup	5	7.9
③ hair_bow	27	42.8
④ hair_side	19	30.2
⑤ hair_fro	25	39.7
⑥ hair_back	13	20.6

表 1 を確認した結果、今回は、30%以上の回答者が異常と判断したケースを基準とすることとし、本方式で排除すべき異常画像は異常①③④⑤にすることとした。

3.2.2 学習用データの準備

3.2.1 の結果を踏まえ、異常画像をチェックする時、1つの画像に複数の異常特徴が存在する場合を前提としてマルチラベル画像分類を導入することとした。

まず、参考文献[2]に示した顔画像生成システムが生成した顔画像 1,000 枚を対象として、3.2.1 の判断基準を踏まえて手作業によりアノテーション作業を行った。

データセットは訓練、検証、テスト用としてそれぞれ 720 枚、180 枚、100 枚を割り当て、画像サイズは 128×128 ピクセルに設定した。各画像には 4 つのラベルが付与され、正常は「0」、異常は「1」で表される。

3.2.3 モデルの構築

事前評価と同様、モデル Keras の Sequential クラスに基づいて構築されており、初めの畳み込み層には 16 個のフィルター、3x3 のカーネルサイズ、ReLU 活性化関数を採用し、入力形状は(ROWS, COLS, CHANNELS)と定義されている。続く層では、32 および 64 のフィルターを持つ追加の畳み込み層と最大プーリング層が配置され、特徴抽出能力を高める。ネットワークは Flatten 層によって平滑化され、512 ユニットを持つ Dense 層が続き、最終層は 4 ユニットの Dense 層で、sigmoid 活性化関数を使用している。モデルには Adam optimizer を使用し、損失関数としては binary_crossentropy が選ばれ、評価指標としては Accuracy が用いられている。

3.2.4 モデルの学習

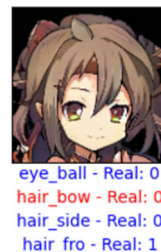
トレーニングは、訓練期間：30 エポック、バッチサイズ：16 で実施した。また、最適化アルゴリズムは、学習率が 5e-5 の RMSprop を採用した。

4. 評価

テストデータ 100 枚を用いて、提案方式の評価を行った。

4.1 画像判別結果

図 4 にテストデータの判別結果の例について示す。今回、正常画像は「0」、異常画像は「1」と判定され、正解は青文字、不正解は赤文字で表示する。



©2014 CloverLab.,Inc.

図 4 テスト画像の判別結果例

図 4 では、その例の結果は、異常⑤の hair_fro は正常に検出されたが、異常③の hair_bow の検出には失敗した。

4.2 正解率

正解率の評価結果を表 2 に示す。また、比較用に同じデータセットをマルチクラス分類にて分類した結果を示す。

本研究の最終目的は異常画像を排除であるため、提案方式において、生成した画像に複数の異常が存在する場合は、1 つの異常を正確的に検出できた場合は、正解と判断し、拡張正解率として整理することにした。

表 2 モデル予測結果

方式	正解率	拡張正解率
マルチクラス	75	75
提案方式	83	92

5. 考察

表 2 を確認した結果、提案方式に基づくマルチラベル画像分類は従来方式であるマルチクラス画像分類と比較して良好な成果を得られることを確認できた。

6. おわりに

本研究では、StyleGAN2 で自動生成したゲームキャラクターの顔画像に対して、機械学習を用いて生成した顔画像を正常画像と異常画像を分類する方式を提案した。具体的には、生成した顔画像に対して異常特徴の抽出に基づく判断基準を設定し、マルチラベル画像分類を用いてモデル構築を行った。評価実験の結果、92%の正解率が得られ、その有効性について確認することができた。今後の課題としては、顔画像自動生成システムへの適用、顔画像自動生成システムの高度化に向けた潜在方向の探索を予定している。

参考文献

- [1] T. Karras, S. Laine, et al.: "Analyzing and Improving the Image Quality of StyleGAN", arXiv:1912.04958 (2019).
- [2] Hironobu Abe and Qing Chen: "An automatic face image generation system with pose and facial expression difference for game characters based on StyleGAN2", Proceedings of the 8th IIEEJ International Conference on Image Electronics and Visual Computing (IEVC2024) CTP-02, pp.1-4 (2024).
- [3] TensorFlow Core, "Image classification", <https://www.tensorflow.org/tutorials/images/classification>.
- [4] K. Prokofiev, V. Sovrasov: "Combining Metric Learning and Attention Heads For Accurate and Efficient Multilabel Image Classification", arXiv:2209.06585 (2022).