

<目的>

- ・ 深層学習に関する基礎知識を学ぶ。
- ・ 実際にAIと画像に関する先行研究をもとにAIがどれくらい違和感なく画像を生成できるか知る。

<データセット>

- ・ <先行研究の手法>で紹介する先行研究が使用したデータセットを図1に示す。
- ・ FFHQ-Aging : 年齢変換アルゴリズムや他の資格タスクのベンチマーク用に設計された人間の顔のデータセット

元のFFHQデータセットの70,000枚の画像には以下の情報が含まれている

- ・ 性別情報
- ・ 年齢層情報
- ・ 顔の傾き
- ・ メガネの種類
- ・ 目の遮断スコア
- ・ フルセマンティックマップ

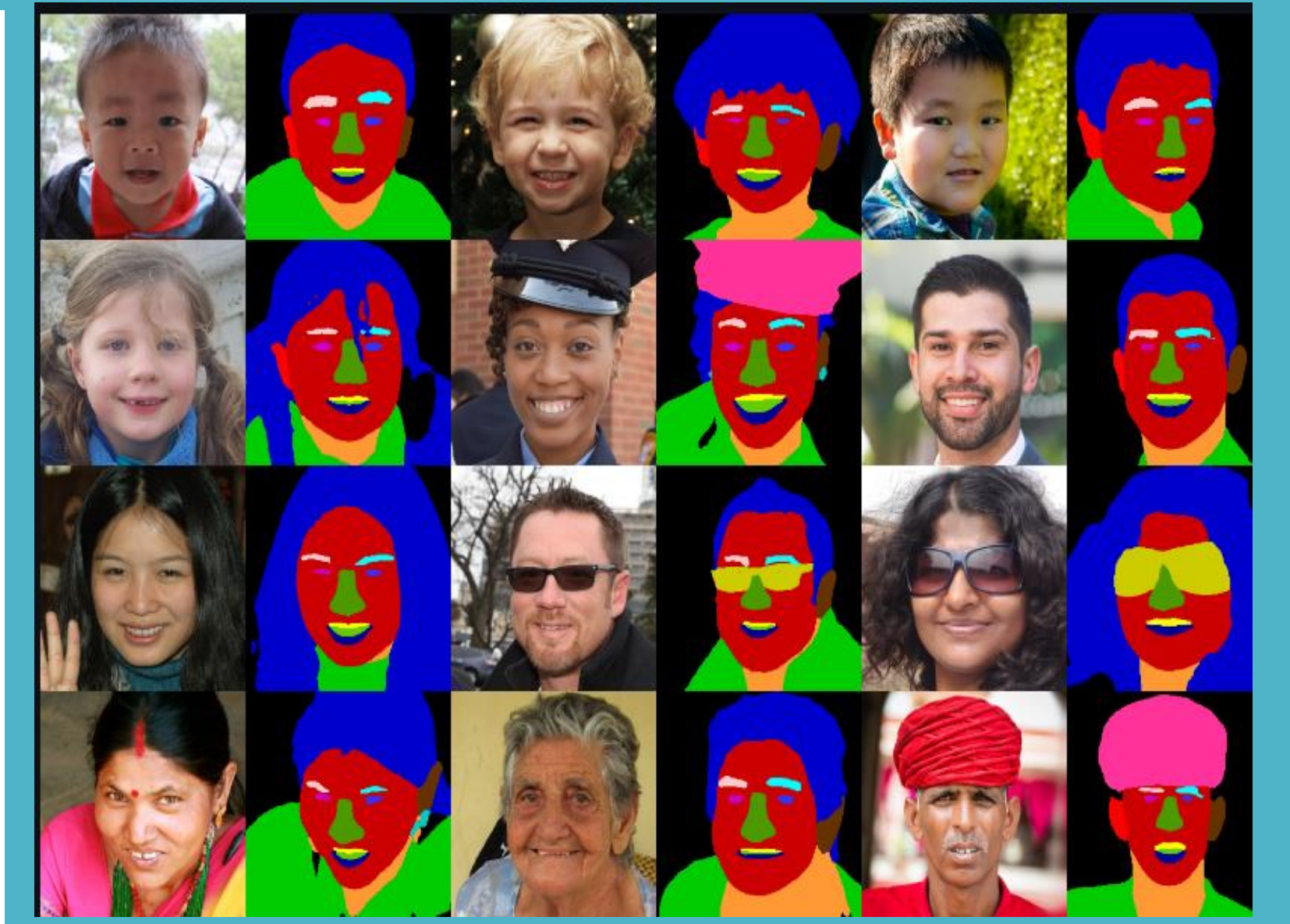


図1.FFHQ-Aging

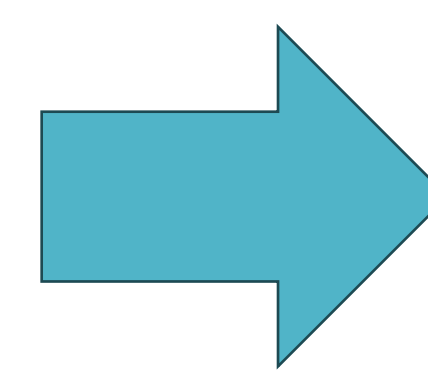
<先行研究の手法>

- ・ 先行研究のgithubで提供されているプログラムを以下の手順で実行して、動画とクラスごとに画像を出力した「Lifespan Age Transformation Synthesis」

米スタンフォード大学、米Adobe Researchによる研究チームが開発した人の顔が過去どのような顔であったか、また未来ではどのような顔になるのかを予測する

基礎知識

- Discriminator(識別モデル)・・・本物かどうかを正しく判定する
- Generator(生成モデル)・・・Discriminatorに偽物と判定されない本物に近いデータを生成する



交互に学習を行うことで本物に近いデータを生成できるGeneratorを作る

アルゴリズム

- ・ 先行研究の論文でアルゴリズムの説明用の図を、図2に示す。

<推論>

- ・ 6つの異なる年齢グループでおおよその継続的な老化を近似することでマルチドメインの画像から画像への変換としての年齢表現を実現し、定式化する
- ・ Generatorは年齢とアイデンティティの間のもつれを解き入力画像からアイデンティティ特徴をエンコードする
- ・ ターゲット年齢クラスが与えられマッピングネットワークを使用してそのターゲットクラスの潜在的なhコードを生成する
- ・ 最後に、デコーダはStyleGAN2の変調量み込みを使用して、潜在的なhコードとアイデンティティ特徴を組み合わせ、連続的なエイジング効果を達成する
- ・ このプロセスでは、隣接する2つの目標年齢クラスの潜在的なhコードを補間することで、画像の年齢を滑らかに変化させることができる

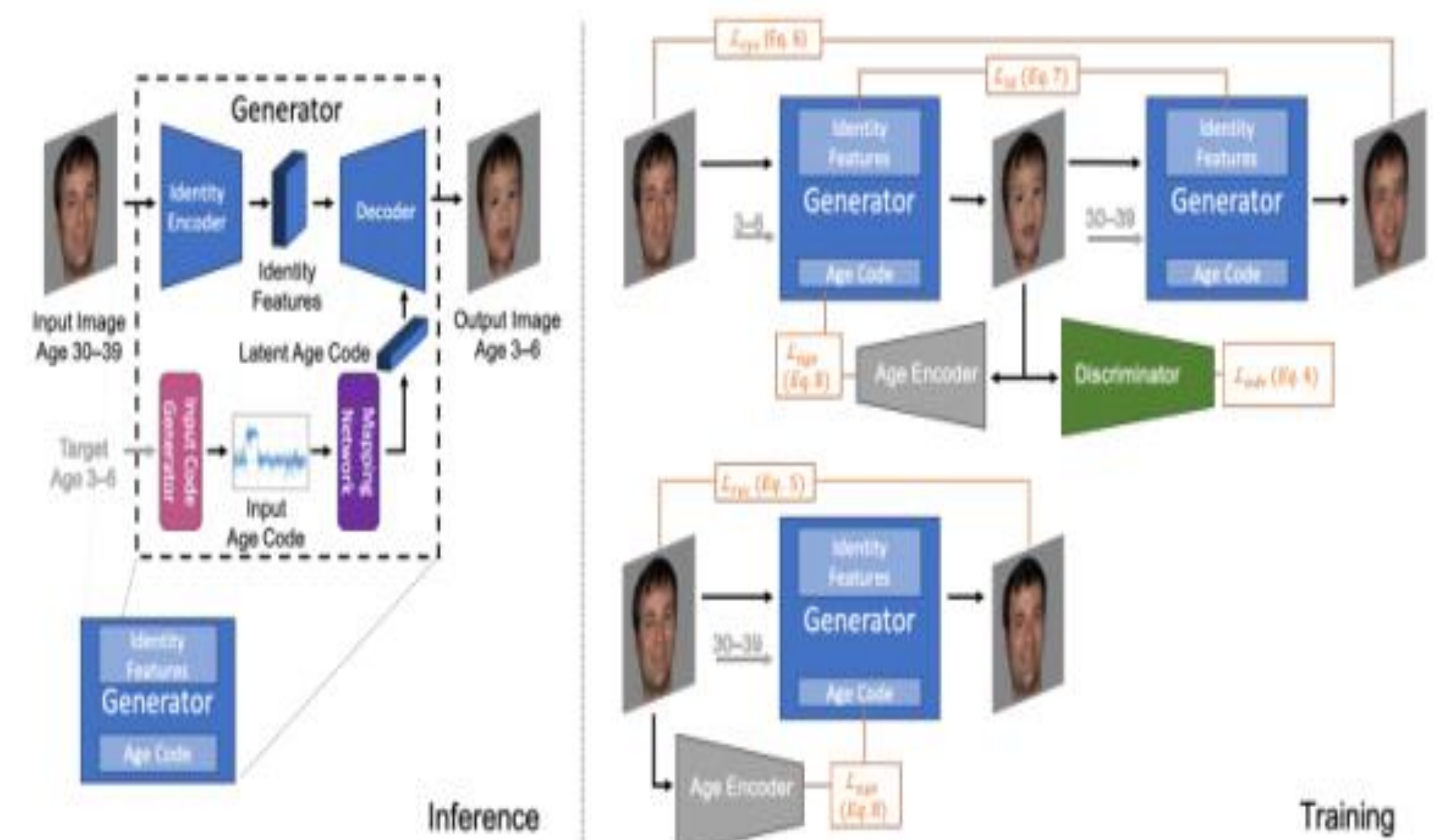


図2.アルゴリズムの概要図

<学習>

- ・ 最初の順方向パスで、次以下の手順でネットワークがトレーニングされる
- ・ 画像を選択し、それを別の年齢クラスに渡して識別器に渡し、結果を現実的に見せて画像と年齢空間の間の接続を強制する
- ・ Age Encoderを使用して生成された画像を潜在的なエッジ空間にマッピングし、間のL1距離を最小化する
- ・ 年齢エンコーダの出力と2番目のフォワードパスのターゲット年齢潜在コード。グラウンドトゥルースの年齢クラスを入力として使用して、変換された画像から元の画像を再生成する
- ・ 3回目の前方パスでは、グラウンドトゥルースhクラスターとともにジェネレータに、画像を供給することで画像を自動エンコードし、ピクセルの差にL1損失でペナルティを課す
- ・ 年齢エンコーダを使って、元の画像を潜在的な年齢空間にマッピングし、学習済みの潜在年齢空間内で補完することにより、連続的な年齢変換を行い、結果を出力する

<結果・考察>

- ・ 年齢変化させた出力結果を図3に示す
- ・ 本手法では顔のしわなどテクスチャの予測だけではなく、頭の形状の予測も行っている
- ・ 極端なポーズやメガネなどの画像には限界がある
- ・ 髪の色や髪型が変化しない

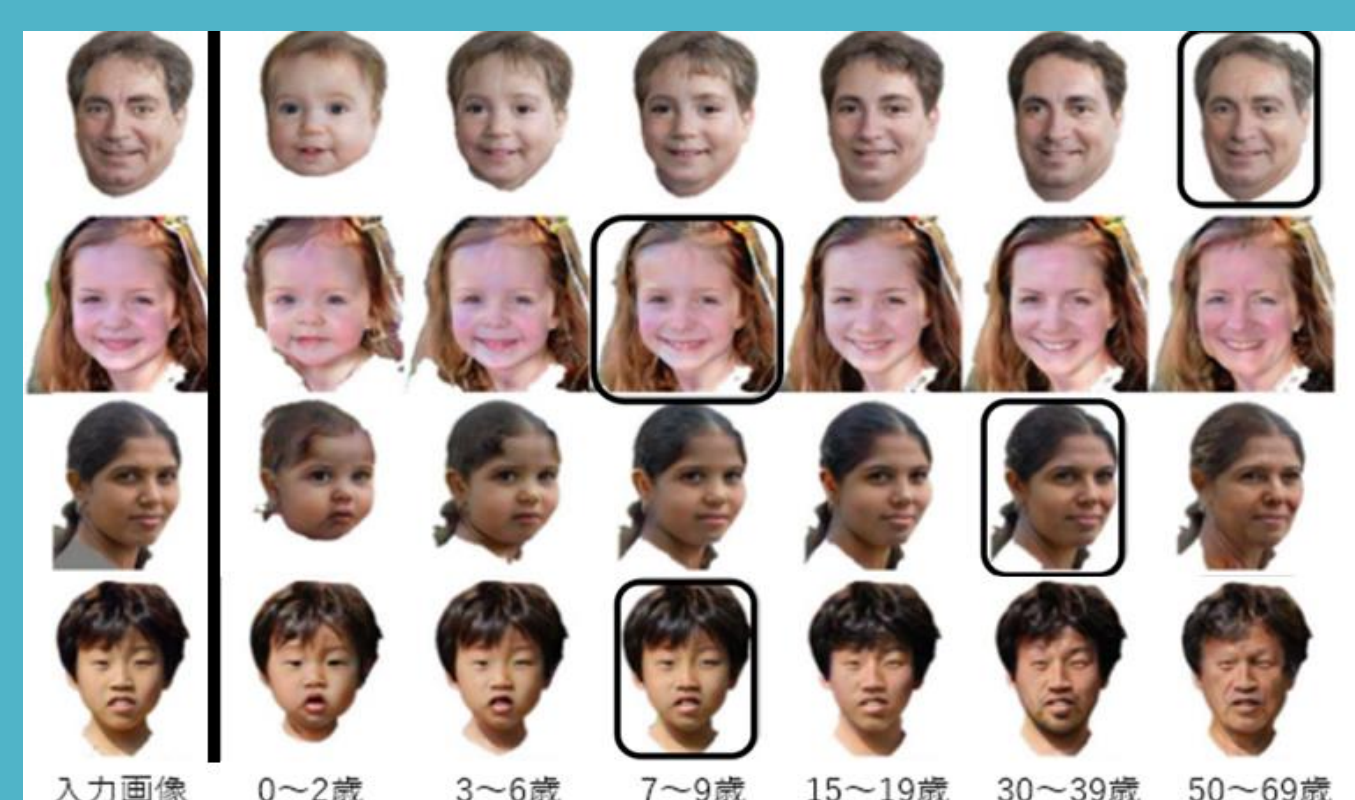


図3.出力結果(例)

<課題・今後について>

- ・ 深層学習の知識不足、計画不足のためにgithub上のプログラムの追試実験しかなかった。
- ・ 先行研究の内容を取り入れつつ一からプログラムを組んでみたかった

参考文献

[1] Roy Or-EI, Soumyadip Sengupta, Ohad Fried, Eli Shechtman, Ira Kemelmacher-Shlizerman : "Lifespan Age Transformation Synthesis" (2020), "FFHQ-Aging". (参照 <https://github.com/royorel/FFHQ-Aging-Dataset>)