

距離場を用いた敵対的生成ネットワークによる植物の3次元形状補間

三浦孝理[†] 清水郁子[†]

[†] 東京農工大学 情報工学専攻

1 はじめに

近年、植物学や農学の分野において情報工学の技術が利用されるようになってきた。植物の3次元形状を取得することで、植物の形状に関する知見を得たりシミュレーションにより農作物の生産効率について検討することが可能になる。3次元形状を得るためには3次元スキャナを用いて複数の方向からデータを取得して統合されることが多いが、植物の葉や枝によってどのような角度で計測したとしても遮蔽された部分が出てしまうことから、植物の3次元形状の欠損を補間することが試みられている [1]。

植物の3次元形状の欠損補間のための従来手法 [1] は、敵対的生成ネットワーク (Generative Adversarial Networks: GAN) によるボクセルベースの手法であるが、十分な精度が得られていない。そこで本手法では、入力を最も近い物体表面への距離の情報を持つ距離場 (Distance Field: DF) を生成して生成器へ入力して学習することで精度向上を試みる。また、従来手法 [1] での識別器は正解データであるか偽物のデータであるかを識別を行っているが、識別機をオートエンコーダとして実装することでより細かい比較を実現することを試みる。

2 提案手法

提案手法のGANのネットワーク構造は図1に示す通りである。欠損のある植物の葉のボクセルデータからDFを生成し、生成器へ入力し、生成器は欠損を補間したDFを出力する。生成器から出力されたDFをボクセルに復元し、識別機へ入力し、学習の偽データとして学習に用いる。識別機へ入力する欠損のない本物データの学習もボクセルデータを入力する。本手法の特徴は従来手法 [1] に対し、生成器にボクセルデータではなくDFを入力すること、識別機にもオートエンコーダを用いることである。

従来の3次元形状モデルの欠損補間はボクセルベースのGANが用いられることが多かったが、物体の存在していないボクセルの割合が多く、学習の際に情報量が少ないという問題がある。そこで提案手法では入力のボクセルデータから最も近い物体表面への距離の

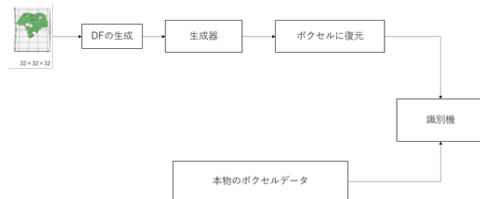


図 1: モデルの全体構造

情報を持たせたDFを生成し、これをGANへの入力とする。

GANの生成器は小針らの手法 [1] と同様に図2のようにデノイジングオートエンコーダを用いる。デノイジングオートエンコーダは与える教師データと入力のデータを異なるデータにすれば入力データから教師データへの変換を学習するようになる。提案手法では入力を欠損のある植物の次元モデルから生成したDF、教師データを欠損のない植物の3次元モデルから生成したDFとすることで、欠損のあるモデルから欠損のないモデルへの変換を学習し、生成器による欠損補間を行う。図2の黄破線の部分をU-Netにおけるスキップ接続とすることでオートエンコーダのエンコーダ部分での入力の形状特徴をデコーダ部分で復元する際に利用することが出来るため、入力データの形状をより再現することが出来る。

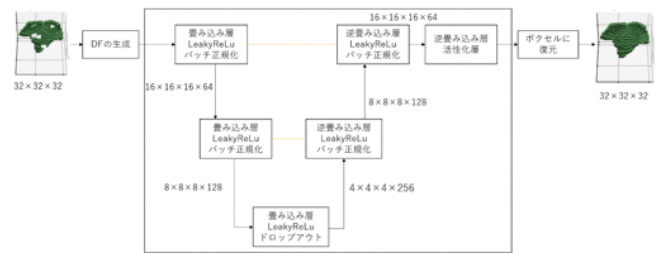


図 2: generator

GANの識別器は一般に入力に対してそれが本物のデータか、生成器が作った偽物のデータであるかの判別しか行わないことが多く、従来手法 [1] でも同様である。しかし、ボクセルデータは物体の存在するボクセルに対して物体の存在しないボクセルの割合が多いので、識別機が本物であるか偽物であるかの判別しか行わない場合、多少無駄に物体の存在するボクセルを増やしたり、物体の存在しないボクセルを作ってしまうてもそれを誤差として認識してしまい、うまく識別ができない。そこで提案手法ではGANの識別機を

3D plant shape inpainting using generative adversarial networks using distance field

Takamichi MIURA[†] and Ikuko SHIMIZU[†]

[†] Faculty of Engineering, Tokyo University of Agriculture and Technology, 184-8588, Tokyo, Japan
s216247w@st.go.tuat.ac.jp

オートエンコーダとして実装し、より詳細に欠損が補間出来ているかを識別する。識別機のネットワークは生成器と同様のネットワーク構造 (図2) である。

識別機の損失関数 L_D は教師データとの平均絶対誤差 MAE であり、通常のオートエンコーダと同様に識別機が入力データから得られた出力が教師データにどれだけ近いかを判別する。生成器は識別と組み合わせで学習する。生成器の損失関数 L_G は L_D と L_{mse} の加重和である。

$$L_D = \frac{1}{n} \sum_{m=1}^n \left[\frac{1}{T^3} \sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \left| \tilde{x}_{m,i,j,k} - x_{m,i,j,k} \right| \right]$$

$$L_{mse} = \frac{1}{n} \sum_{m=1}^n \left[\frac{1}{T^3} \sum_{i=1}^T \sum_{j=1}^T \sum_{k=1}^T \left\{ \tilde{x}_{m,i,j,k} - x_{m,i,j,k} \right\}^2 \right]$$

$$L_G = \alpha_1 L_{mse} + \alpha_2 L_D \quad (1)$$

ここで、 T はボクセル空間の1辺あたりのサイズである。教師データとの平均二乗誤差 L_{mse} は生成器のデノイジングオートエンコーダが入力データから教師データへの変換に用いられるものであり、 L_G は L_{mse} と L_D (生成器の学習では識別機に誤認させるために本物のデータが正解のデータとなる) に重み α_1, α_2 をかけた和とする。

3 実験

3.1 学習

学習に用いるデータセットは植物を3次元スキャナで計測し、その植物の葉の部分抽出したモデルを用いる。3次元スキャナで得られた点群データをボクセル化し、9枚の葉のモデルを得る。得られた9枚の葉のボクセルデータに回転、移動、反転を適用しデータの拡張を行う。拡張したデータのボクセル空間のランダムな領域を削除して欠損データを生成する。以上の操作により欠損のあるデータと欠損のない教師データの288組を生成した。

識別機を学習する際の正解のデータの学習は欠損のないデータのボクセルデータで行う。実験では、ボクセル空間のサイズ T は32とし、 $\alpha_1 = 0.999, \alpha_2 = 0.001$ とした。学習後のテストデータは学習の最に使用していないデータを用いる。

3.2 結果

実験結果を図3, 図4に示す。図3は小領域に対する実験結果、図4は大領域の欠損に対する実験結果である。両方の図で上の左から欠損のないデータ、欠損を与えたデータ、下の左からDFを入力とし従来手法 [1] と同様の識別機出の実験結果、入力をDFとし識別機をオートエンコーダにした時の実験結果である。

図3, 図4よりDFを導入したことにより、生じてしまった欠損を従来手法よりも補間出来ていることがわかる。特に通常のボクセルデータでは大領域の欠損をほとんど補間できていなかったが、DFを導入することによって大きく改善した。しかし、識別機をオートエンコーダとして実装した結果は欠損の補間もあま

り出来ず、元の形状から大きく崩れるものとなってしまった。

学習の際に損失関数の重みを工夫するか、元のデータに何も無いところに物体を作ってしまった場合や、物体のある所から物体を欠損させてしまった場合などにペナルティを与えることで元の形状をより再現できるようになるのではないかと考えられる。

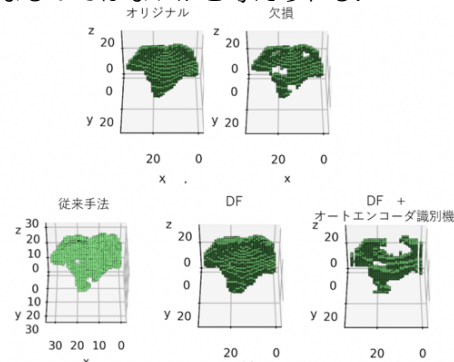


図3: 小領域の欠損に対するオートエンコーダ識別機の実験結果

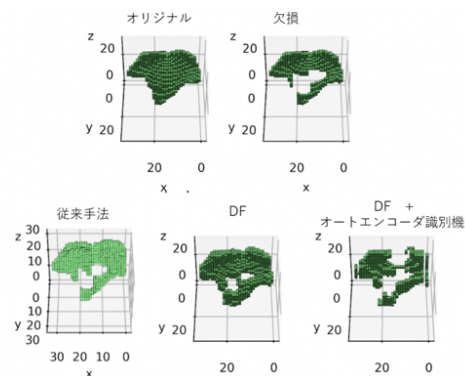


図4: 大領域の欠損に対するオートエンコーダ識別機の実験結果

4 まとめと課題

本稿では、植物の3次元形状の欠損を補間するために、DFやオートエンコーダ識別器を導入したGANによって欠損を補間する手法を提案した。

実験ではDFの導入による欠損補間における有効性を示した。特に大領域の欠損に対しての有効な結果を示した。元の形状を再現するためにGANの識別器をオートエンコーダにした場合には、元の形状が崩れてしまい、良い効果は確認されなかった。今後は、損失関数を調整し、ペナルティなどを与えることによって形状を再現できるような改善を行う。

参考文献

[1] 小針和也, 諸橋陽王, 横山圭, 清水郁子, “植物の3次元形状モデリングのための敵対的生成ネットワークによる形状補間,” 映像情報メディア学会技術報告 43巻26号 pp55-58(2019)