

網膜静脈分岐閉塞症を生じた網膜 OCT 画像に基づき視力を推定する DNN の開発

1. 研究背景・目的

網膜静脈分岐閉塞症 (Branch Retinal Vein Occlusion, 以下 BRVO) は網膜内の静脈が枝分かれしている部分で血栓が形成され、行き場を失った血液が血管から溢れ出る疾患であり視力低下を引き起こす。この BRVO による出血自体は時間とともに消失する。出血消失後の視力回復の度合いは眼底の中心部に位置する黄斑への影響度 (重度の場合、黄斑に浮腫が形成される) に依存する。現行の治療法としては、血管内皮増殖因子を阻害する薬剤を眼の硝子体に直接注射する方法がある。しかし注射投与は高価かつ必ずしも注射投与後に視力が回復するとは限らない。さらには現状、注射の判断は各眼科医が経験に基づいて独自に行うため、統一的な判断基準が存在しないことも問題である。

これらの問題を解決する方法として、注射前に注射投与後の視力回復の程度を推定するニューラルネットを開発する。これまでに取得した患者の網膜 OCT 画像と注射投与前後の視力変化の関係を学習することで、投与による視力の改善量を正確に推定できれば眼科医の投与判断の助けになり、患者の予後の QOL 向上、治療コストの低下につながる。

2. 本研究で使用するデータ及び手法

本研究では二つの検証を実施する。一つは、学習データに機材 1 で撮影した画像のみを使用し、テストデータに両方の機材で撮影された画像を用いて精度を比較する (以下、検証 1)。もう一つは、学習データとテストデータ両方に二つの機材で撮影された画像を使用して精度を比較する (以下、検証 2)。これらの結果により、同じ OCT 画像でも異なる機材で撮影された画像を同時に学習させる有用性の検討を試みる。

2.1. 使用するデータ

2.1.1. 検証 1

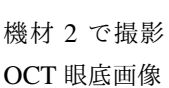
使用するデータは 85 人の BRVO 患者から 2 つの機材で撮影した 102 枚の画像である。Fig. 1 に機材 1 で撮影された画像、Fig. 2 に機材 2 で撮影された画像を示す。ニューラルネットの学習に使用するものは、機材 1 で撮影された画像 68 枚、テスト画像として使用するものは、機材 1, 2 共に 17 枚である。本研究では一般的な画像認識の研究で使用されるデータ数より少ないため、一枚の画像に種々の加工 (わずかに回転やスライド) を施しデータを 1 枚あたり 30 枚に水増しする。最終的にニューラルネットの学習に使用する画像は 2040 枚 (68×30) である。

2.1.2. 検証 2

使用するデータは 85 人の BRVO 患者から 2 つの機材で撮影した 144 枚の画像である。ニューラルネットの学習に使用するものは、機材 1, 2 で撮影された画像 116 枚 (各 58 枚)、テスト画像として使用するものは、機材 1, 2 共に 14 枚である。検証 1 同様、画像のオーグメン



Fig. 1 機材 1 で撮影された OCT 眼底画像



テーションを実施しているので、最終的に学習に使用する画像枚数は 3480 枚である (116×30)。検証 1, 2 のデータ枚数の決定の方法は、appendix を参照されたい。

2.2. 転移学習

少ない画像データに対し、少ない計算時間で高精度のモデルを構築する「転移学習」を使用する。すでに大量の画像データから特徴を学習してあるニューラルネットを使用する手法であり、医療画像などのデータが少ないものに対して良く使われる手法である。以下に転移学習を実施する際の例を示す。なお本研究ではニューラルネットのモデルとして GoogleNet の inception-V3 を使用している。

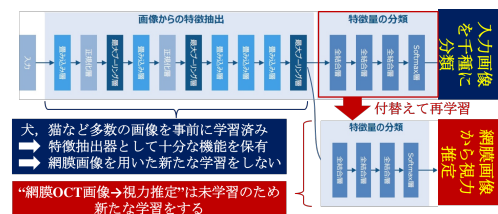


Fig. 3 転移学習のモデル

3. 結果および考察

3.1. 検証 1

検証 1 では、バッチサイズを 40、エポック数を 750、過学習やデータの振り分け方の違いによる偏りをなくするため 5 分割差検証を実施する。最適化手法には Adam を使用し、学習率は 0.0002 でそれ以外はデフォルトの値を使用している。NVIDIA 製 GPU の GTX1080Ti を使用して計算を実施する。

横軸に推定値と真値との誤差許容範囲、縦軸に許容範囲内に収まるテストデータの割合としてプロットしたものを Fig. 4 に示す。橙線が機材 1 で撮影された画像をテストデータ、青線が機材 2 で撮影された画像をテストデータに使用した時の結果である。なお表示してあるものは 5 分割差検証した各モデルにおける予測の平均値である。学習データと同じ機材で撮影された画像をテストデータに使用した方が概ね推定精度が良いことがわかる。誤差許容範囲が 0.2 において、機材 1 の画像をテス

トデータに使用した場合（橙線）60 %程度であるが、機材 2 で撮影された画像をテストデータとして使用した場合（青線）、精度は 50 %程度である。多少精度は落ちているが、異なる機材で撮影した画像に関わらず Fig. 4 の結果が得られたのは、OCT 眼底画像から視力推定するタスクにおいて転移学習を使用し、事前学習された画像の特徴を利用する有用性を示していると言える。

3.2. 検証 2

エポック数を 250 に減らした点以外は検証 1 と同条件である。横軸に推定値と真値との誤差許容範囲、縦軸に許容範囲内に収まるテストデータの割合としてプロットしたものを Fig. 5 に示す。橙線が機材 1 で撮影された画像をテストデータ、青線が機材 2 で撮影された画像をテストデータに使用した時の結果である。なお表示してあるものは 5 分割交差検証した各モデルにおける予測の平均値である。傾向としては、テストデータに機材 2 で撮影された画像を使用の方が推定精度がよいことがわかる。これは Fig. 2 に示すように、眼底の情報がグレースケール画像で表示した時に、多く含まれているからと考える。また、誤差許容範囲が LogMAR 視力 0.2 において、機材 1 で撮影された画像をテストデータとして使用した場合（橙線）の推定精度は 60 %程度であるが、機材 2 で撮影された画像をテストデータとして使用した場合（青線）の推定精度は 70 %程度である。検証 1 よりも推定精度が上昇している理由としては、テスト枚数の増加が考えられる。しかし、機材 1 での画像データ数に注目すると、検証 1 よりも学習データが少ないにも関わらず、同程度の精度が出ているということは、機材 2 で撮影した画像からも機材 1 の推定に役立つ有効な特徴を抽出していると考えられる。

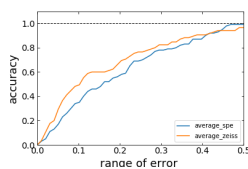


Fig. 4 学習データに機材 1 のみ使用したモデルでの推定精度

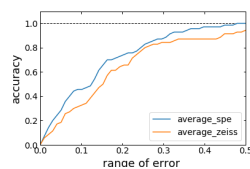


Fig. 5 学習データに機材 1,2 両方使用したモデルでの推定精度

横軸に LogMAR 視力、縦軸に推定誤差として表示したものを Fig. 6 に示す。LogMAR 視力 0.2 から離れるほど、視力推定精度が悪くなる傾向がわかる。Fig. 7 から LogMAR 視力 0.2 や 0 のテストデータが多いことから、Fig. 6 のような結果の理由は、データ数が多い LogMAR 視力ほどそれに伴い、推定精度も向上することが原因と考えられる。LogMAR 視力のレンジを変更しても一定の推定精度を出力するには全ての LogMAR 視力において、偏りなくデータが存在する必要がある。

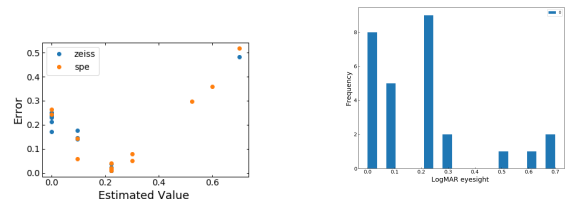


Fig. 6 LogMAR 視力と誤差の関係

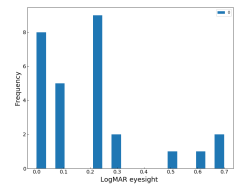


Fig. 7 テストデータに使用した LogMAR 視力のヒストグラム

4. 今後の予定

さらなる精度の向上に向けて、機材 1, 2 で撮影された画像に加えて治療前の視力を入力に加えるマルチモーダル学習を試みる。また OCT 画像の重要な部位をセグメンテーション等で抽出し、視力推定に特に役立つ領域に対して、ニューラルネット内で重み付けを行うことで精度の向上を図る。

5. Appendix

検証 1 と検証 2 に使用した画像枚数の詳細を以下では示す。

5.1. 検証 1

機材 1 で撮影された画像が 85 枚である。5 分割交差検証を実施するため、テストデータは各検証において 17 枚になる。よって学習データは $85-17=68$ 枚になる。

5.2. 検証 2

機材 2 で撮影された画像は 72 枚である、両機材で撮影された画像の枚数を揃えるために、機材 1 で撮影された画像も 72 枚使用する。検証 1 の時と分割比率を同じにすると、各機材において学習データは 58 枚、テストデータが 14 枚になる。