

Preferred Networks インターン選考 2018 コーディング課題 Chainer 活用分野

- Chainer を用いた画像処理を行う課題である。
- python を用いて回答を作成し、ライブラリとしては標準ライブラリに加えて chainer, numpy, opencv を使用すること。
- chainer は最新版 (4.0.0) を用いること
- opencv はデフォルトが BGR の実装と RGB の実装とがあるが、BGR デフォルトを用いること
- 課題には自分だけで取り組んでください。この課題を他の応募者を含めた他人と共有・相談することを禁止します。漏洩の証拠が見つかった場合、その応募者は失格となります。ある応募者が別の応募者に回答をコピーさせた場合、双方の応募者が失格となります。
- 想定所要時間は最大 2 日です。全課題が解けていなくても提出は可能ですので、学業に支障の無い範囲で取り組んで下さい。

提出物

- 問題 1 ソースコード、学習済みモデル (serializers.save_npz で保存)。使用した Python のバージョンは明記してください (例えば 2.7, 3.6 など)
- 問題 2 ソースコード、学習済みモデル (serializers.save_npz で保存)、(3) については考察を記載した (pdf/word/txt)

評価基準

提出物の評価にあたって以下のような要素を考慮します。(必須では無いので、時間がない場合はこれらを満たしていなくても大丈夫です)

- ソースコードが他人が読んで理解しやすいものになっていること。
- ソースコードが正しいものであることを検証するためにある程度の単体テストや検証のためのコードが書かれていること。
- 提出物を見て追試がしやすい形になっていること。
- レポートが要点についてわかりやすくまとまっていること。

提出方法

上記の提出物を単一のパスワード無し zip ファイルにまとめ、[こちらの](#)専用フォームより応募してください。締切は日本時間 2018 年 5 月 14 日 (月) 23:59 です。

問い合わせ

問題文に関して質問がある場合は intern2018@preferred.jp までご連絡ください。問題文に訂正が行われた場合は応募者全員にアナウンスいたします。なお、アプローチや解法に関する問い合わせにはお答えできません。

問題 1

画像に対する学習を行う時にフィルター等で変換を行ってから学習する時があるが、フィルター自体も学習で作る事ができる。学習ベースのフィルターはデータがあれば複雑なフィルタも作れる点や、それ自体が微分可能である点がメリットとしてある。以下では、RGB 色空間から Lab 色空間への変換を行うフィルタの学習を課題とする。

- (1) 1D コンボリューション, ReLU 関数を用いた 3 層のニューラルネットワークモデルを定義し RGB \rightarrow Lab 色空間への写像を行うネットワークをトレーニングせよ。OpenCV に RGB2Lab 色変換の関数があるので、それを用いて変換前変換後のデータのペアを出力する DataSet クラスを定義し、Trainer を用いてトレーニングする事。
- (2) 学習中のモデルに対して変換精度の評価を行いたい。RGB の色空間全域について目標値との最小 2 乗誤差を計算し、平均誤差と最悪誤差を出力する evaluator を定義せよ。また、実際に (1) で学習したモデルをロードして training する際、一定 iteration 事に evaluation を実行出来るように設定し、動作を確認せよ。(RGB 全空間でチェックを行うと実行時間がかかりすぎる場合は適宜間引いて実行してもよい)
- (3) Lab \rightarrow RGB 色空間への逆変換を行うネットワークについても同様に学習し evaluator を用いてチェックせよ。

問題 2

ディープラーニングを用いた画像認識の字用例として画像分類やオブジェクトのディテクション、セグメンテーション等が考えられる。Chainer では VGG, ResNet 等の学習済みモデルが使用可能であるが、以下ではそれを特定の問題設定に応用する事を課題とする。

- (1) Data フォルダに与えられた ImageNet 由来の画像データ (imagenet_data) と非 ImageNet 由来の画像データ (anim_face_data) を識別するネットワークを作成したい。いずれかの学習済みモデルをロードし、全結合層や出力層をファインチューニングすることで、ネットワークをトレーニングせよ。トレーニング用のデータは必要に応じてリサイズ、クロップせよ。
- (2) ImageNet 由来の画像データを background, 非 ImageNet 由来の画像データを foreground として、foreground を切り出すネットワー

クを作成したい。まず、backgroundとして用いる imagenet_data の画像を anim_face_data の倍のサイズに拡大し、ランダムな位置に anim_face_data を合成した画像と、合成した位置を示すマスクを出力するデータセットクラスを作成する。次に (1) のようにトレーニング済みモデルをロードし、その中間層（例えば全結合層の直前の層）の特徴量を入力としてマスク画像を予測するようなネットワークを定義し、それをトレーニングせよ。（この問題では合成画像データが一定サイズの矩形で合成されていることを既知としてもよい、しなくてもよい）

- (3) 上記 (2) のような合成データを用いたトレーニングは目的に応じたモデルを作るうえで重要であるが、単に検出対象画像を合成するだけでは十分な検出性能が出ないことが考えられる。例えば上記 (2) のモデルが実データに対して精度が十分に出ない場合、理由と対策としてどのような手法が考えられるか、単純にデータを増やす以外にいくつか挙げて期待される効果を説明せよ。