

ディープラーニングを用いた Web サイトデザインの年代推定

プロジェクトマネジメントコース 矢吹研究室 1442104 増田 準

1. 背景

Web サイトのデザインは、時代に合ったものが求められる [2]。スマートフォンの爆発的な普及により、Web サイトは急速に発展を遂げた。Web サイトをデザインするということは、視覚的な良し悪しを求めるだけでなく使いやすさなど様々な要素を含む。その為 Web サイトを閲覧するデバイスによってデザインを変える事などもあり、現代における Web デザインの多様化は著しい。更に、人々の生活に密接に影響していることから、Web サイトに対する研究は学際的に取り組まれている [1]。以上のことから、本研究では時代によって進化する Web デザインの解析を対象とする。

2. 目的

この研究では、年代ごとの Web デザインの変化を解析することを目的とする。デザインとは数値などで表すことができるものではなく、漠然としたものである場合が多い。その為、解析の際は Web サイトに使用されている文字のフォントやその配置、色彩や画像など、ページに映る要素を総合的に判断させることが重要だ。

3. 手法

この研究は以下の手法を用いて行う。

3.1 機械学習

機械学習による画像解析を利用する。この研究における画像解析とは、多数の教師画像を学習させ判別モデルを作成し、別の画像を判別させることで画像の特徴を解析する処理を指す。

3.2 画像の収集

Internet Archive にて閲覧できる Web ページを、Selenium によるスクリーンショットを用いて多数保存する。保存の対象となる Web サイトは 2017 年度版の Fortune Global 500 [3] にリストされた企業のホームページとする。

3.3 画像解析

保存する画像は、Web サイトの公開年を 1996 から 2002、2003 から 2009、2010 から 2017 という

3 世代に分類しタグ付けする。そしてこれらの画像を用いて、TensorFlow による画像解析を行い Web デザインのみで年代を判別可能にする。

4. 想定される成果物

想定される成果物は以下のとおりである。

- Web サイトの年代解析結果及び予測モデル
- Web サイトの年代が判別可能となる教師画像のデータセット
- 機械学習による年代解析用のコード

5. 進捗状況

上記した画像の収集にある手順で教師画像を集めた。スクリーンショットを実行するコードにおいてサイズの指定をしているが、キャプチャのサイズにばらつきが出ることや、Internet Archive を読み込む際の処理落ちによりエラー表示のまま保存されているものが多かった。エラー表示のまま保存された画像を除き、7000 枚強ほどの教師画像を収集した。

6. 今後の計画

機械学習による画像解析用のコードの書き方を習得し、解析を実行する。解析の結果によって教師画像の追加や、リサイズを洗練させるなどして正解率の向上を図る。

参考文献

- [1] 酒巻隆治, 染矢聡, 岡本考司. Web デザインに対する印象と記憶される情報量との関係性分析. 日本デザイン学会研究論文集, Vol. 55, No. 6, pp. 59–66, 2008 年.
- [2] こもりまさあき, 赤間公太郎. Web デザインの新しい教科書. エムディエヌコーポレーション, 改訂新版, 2016 年.
- [3] Fortune Global 500. Fortune global 500 list 2017. <http://fortune.com/global500/list/> (2017.09.15 閲覧).