

# ラスタ形式画像を学習用データとするグラフィックデザイン評価モデルの構築

石山 航平<sup>†</sup> 牛尼 剛聡<sup>††</sup>

<sup>†</sup>九州大学大学院芸術工学府 〒815-8540 福岡県福岡市南区塩原 4-9-1

<sup>††</sup>九州大学芸術工学研究院 〒815-8540 福岡県福岡市南区塩原 4-9-1

E-mail: <sup>†</sup>tshiyama.kouhei.842@s.kyushu-u.ac.jp, <sup>††</sup>ushiyama@design.kyushu-u.ac.jp

**あらまし** 本論文ではグラフィックデザインの自動生成やユーザが生成した画像の自動評価を対象として、アノテーションが行われていないグラフィックデザインのラスタ画像を学習データとして、デザインの適切さを自動評価する機械学習モデルを提案する。グラフィックデザインは様々なデザイン的な観点を考慮して構成する必要があるが、提案モデルでは、レイアウトと配色の観点から与えられたグラフィックデザインを評価する。提案手法では、教師データとなるグラフィックデザインの画像をそのレイアウトと配色の情報を抽出した画像に変換し、それらを用いて敵対的生成ネットワーク (GAN) を学習させる。そこで学習された識別器を利用して、指定された素材を利用して自動的に生成したデザインのレイアウトと配色の評価を行う。提案手法の有効性を評価するために、被験者を利用した主観評価による実験の結果、提案手法の有効性が示された。

**キーワード** GAN, 自動デザイン, 深層学習, 配色, レイアウト

## 1 はじめに

近年、機械学習を利用したコンテンツ生成に注目が集まっている。敵対的生成ネットワーク (GAN) [1] や拡散モデル [2] のような深層学習を用いたコンテンツの自動生成技術は飛躍的な進歩を遂げている。

コンテンツの自動生成が期待される分野には様々なものが存在する。その代表的な分野の一つに、グラフィックデザインがある。現在、広告やポスターなどといったグラフィックデザインの多くが専門的な知識と経験を持ったプロのデザイナーによって制作されている。機械学習技術によってデザインの自動生成を行うことができれば、ユーザの特性や目的に応じて、最適なコンテンツを自動的に提供可能になる等、様々な応用が期待できる。

コンテンツの自動生成を行う代表的な手法の一つに GAN がある。GAN では生成器と識別器という 2 種類のニューラルネットワークから構成される。生成器はノイズからコンテンツを生成するニューラルネットワークである。識別器は入力されたデータが学習データ集合に含まれる本物のデータであるか、生成器で生成された偽物のデータかを見分けるニューラルネットワークである。生成器と識別器が敵対的に学習を行うことによって、生成器が高品質なコンテンツを生成できるようになる。

これまでにも、GAN を利用してグラフィックデザインのレイアウトを生成するいくつかの手法が提案されている [3] [4]。それらは学習データとして、利用するグラフィックデザイン画像に含まれる視覚要素のレイアウト情報が必要である。ここでのレイアウト情報は人手によるアノテーションによって作成されるのが一般的であるが、作業のコストは大きいという問題が



図 1 提案手法の概要

ある。

本論文では、学習用データとして与える画像データは、そこに含まれる視覚要素について配置や色などのデザインパラメータに関する情報が得られることを前提とせず、アノテーションが行われていないラスタ形式の画像データを用いる。これにより、既存の大量の画像データを、そのまま学習データとして利用できるため、様々な種類のグラフィックデザイン生成に適用可能という利点がある。

本論文で提案する手法の概要を図 1 に示す。本論文では対象とするグラフィックデザインの種類としてバナー広告を例題として利用する。具体的には、商品画像やイメージ画像といった素材画像と広告文を入力として、それらの素材の特徴に応じて、広告文の配置や大きさを自動的にデザインし、バナー広告として適切な広告画像を生成する。

本研究による貢献は以下のとおりである。

- グラフィックデザインの構成要素のレイアウトや配色に関するアノテーションを用いず、ラスタ画像のみを学習用データとして利用し、グラフィックデザインの評価機構を提案

した。

- 被験者による評価実験により提案手法の有効性を示した。

## 2 関連研究

これまでもグラフィックデザインの自動レイアウトに関する研究が活発に行われてきた。本章では、それらの研究について述べ、本研究の位置づけを明確にする。

Sreekanth [5] らは、遺伝的アルゴリズムを用いて、バナー広告のレイアウトを生成する手法を提案している。この手法では、バナー広告を構成する要素の重複や、要素間の距離に関する評価指標を定義し、それらを遺伝的アルゴリズムによって最適化することによって、バナー広告のレイアウトを自動的に決定する。

Jianan [3] らは、GAN を利用して、様々なグラフィックデザインのレイアウトの生成を行う手法を提案している。この手法では、各要素を配置するためのパラメータを生成する生成器と配置に関するパラメータ群からレイアウト情報を表すワイヤフレーム画像に変換し、畳み込みニューラルネットワーク [6] によって本物か偽物かを判別する識別器の機構を利用している。

Xinru [4] らは、入力に応じた最適雑誌のレイアウトの自動生成手法を提案している。この手法では、入力素材の特徴を VGG16 [7] や Word2vec [8] によってエンコードし、それらを生成器と識別器に組み込み、生成器のノイズを入力に応じた条件付き分布からサンプリングする。

上記の手法は全て、学習データとするグラフィックに関して、それを構成する画像やテキストの要素の位置やサイズ等のレイアウト情報が利用できることを前提としている。しかし、グラフィックコンテンツを自動生成するために、そのグラフィックコンテンツを構成するレイアウト情報が利用可能であるとは限らない。特にインターネット上に存在する広告画像を学習データに利用しようとしても、それらに関しては、構成要素のレイアウト情報をデータとして利用できない場合が多い。レイアウト情報がわからない結果としてのグラフィックコンテンツを学習データとして利用することができれば、より汎用性が高いグラフィックコンテンツ生成手法となる。本論文では、各要素のレイアウト情報が与えられていないラスタ形式の画像を学習用データとし利用する手法を提案する。また、本手法においてはレイアウトだけでなく配色という観点も取り入れ、複数の観点からのグラフィックデザイン評価を行う。

## 3 提案手法

### 3.1 全体像

グラフィックデザインには様々な種類が存在する。本論文では、例題としてバナー広告を対象とする。本論文では、ユーザが対象とするバナー広告の構成要素として以下を明示的に入力として与えることを前提とする。

- 素材画像

- ロゴ画像
- 広告文
- ボタン画像内容文

本論文では、二つの機構によって構成されるグラフィックデザインの自動生成システムを想定する。一つはデザイン機構 *Design* であり、入力された構成要素を利用してグラフィックデザインの生成を行う機構である。具体的には、素材画像 *im*、テキスト *text*、構成要素の位置、大きさ等のレイアウトに関するパラメータ  $\mathbf{p}^l$ 、配色に関するパラメータ  $\mathbf{p}^c$  に基づいて、画像集合を生成する。

$$Design(im, text, \mathbf{p}^l, \mathbf{p}^c) = (g_1, g_2, \dots) \quad (1)$$

$$\mathbf{p}^l = \begin{pmatrix} \text{素材画像の座標} \\ \text{素材画像の拡大率} \\ \text{ロゴ画像の座標} \\ \text{ロゴ画像の拡大率} \\ \text{広告文の座標} \\ \text{広告文の拡大率} \\ \text{ボタン画像の座標} \end{pmatrix}$$
$$\mathbf{p}^c = \begin{pmatrix} \text{背景色} \\ \text{広告文の色} \\ \text{ボタンの色} \end{pmatrix}$$

もう一つは評価器 *Evaluate* であり、評価器はデザイン機構で生成されたグラフィックデザインの画像集合の要素  $\mathbf{g}$  に対して、デザインの適切さを評価しスコア  $s$  を返す関数として形式化される。

$$s = Evaluate(\mathbf{g}) \quad (2)$$

システムは、評価器が出力したスコアが高いものをバナー画像候補としてユーザに提示する。また、本手法ではグラフィックデザインの重要な要素である「レイアウト」と「配色」という観点から生成を行う。まずレイアウトのみを生成し、その後、生成されたレイアウトを用いて配色の決定をする。

### 3.2 デザイン機構

ユーザが入力したバナー広告画像の構成要素はレイアウト生成段階では、バナー画像上での位置、サイズ、配色決定段階では背景色、前景色の情報をデザインパラメータとして扱う。このとき、バナー広告をデザインするという処理は、それぞれの構成要素に対して適切なデザインパラメータを決定することになる。本手法では、デザインパラメータの決定は、デザイン機構によって行われる。今回はデザイン機構では、乱数によってそれぞれのパラメータを決定してバナー画像候補集合を生成する。

### 3.3 評価器

デザイン機構によって生成されたバナー広告画像候補に対して、評価機構が、バナー広告画像としてのデザインの適切さを評価する。デザインの適切さは様々な観点から考えることがで

きる。本研究では、インターネット上に存在する大量のバナー広告の適切さを、レイアウト、配色の二つの観点から捉え、それぞれの観点からの適切さを評価する評価器を利用する。評価器は、GANを利用して学習させた識別器を用いる。

大量のバナー広告を学習データとして、GANによって学習が行われた識別器の出力はGANの種類によって違いはあるが、そのほとんどが、入力データが本物のバナー広告画像であると判断するための確率や距離といった数値である。本研究ではバナー広告画像はデザイナーがデザインの専門的な知識を利用して作成されていると考える。そこで、デザイン機構が自動的に生成したバナー広告画像が、既存のバナー広告画像と区別することができなければ、その候補はバナー広告画像として適切な性質を有する可能性が高いと考える。GANを利用して識別器と生成器の学習を繰り返すことによって、学習済みの識別器は入力画像に対して、潜在的な「バナー広告のデザインとしての妥当性らしさ」を評価可能となる

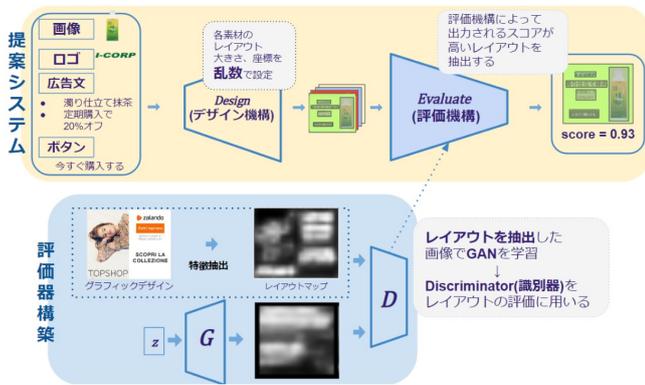


図2 提案手法の全体像

## 4 評価機構の構築

本手法では、GANによる学習済みの識別器を、生成された画像がバナー広告画像としての妥当性を評価する評価器として利用する。グラフィックデザインの良し悪しを決める構成要素として、写真や画像などのビジュアル素材、文字、色、レイアウト、線・飾りなど様々な観点が存在する[8]。本研究においてはレイアウトと色のみ注目し、それらについてデザイン的に適切かどうかの評価を自動的に行うことを目指す。それぞれの観点からのデザイン的な妥当性を学習するため、学習データに前処理を加えることによって、「レイアウト」、「配色」という2つのデザインの観点レイアウト情報について評価を行う評価器を構築する。まずデザイン機構によって生成されたものの中でレイアウト評価器の出力の値が高いレイアウトを抽出し、そのレイアウトを適用し、デザイン機構によって色情報を決定し、生成されたものの中で配色評価器の出力の値が高いデザイン画像を抽出する。

画像の生成を目的とするGANとして様々な手法が提案されているが、本研究では、GANのモデルとしてDCGAN[9]とPGGAN(Progressive Growing of GAN)[10]を採用する。

DCGANは識別器にネットワークを全結合層のみではなく、畳み込み処理を加え、生成器のネットワークには畳み込みの逆の処理である、転置畳み込みを加えたGANのモデルである。DCGANの識別器の出力は $y$ は入力されたデータが学習データ集合に含まれる確率を表しているため、評価器として用いる際、 $y$ をそのままレイアウトスコアや配色スコアとして用いる。DCGANで構築された評価器のスコアは大きいほどデザインとして妥当だとみなされる。

PGGANは識別器に畳み込み層、生成器に転置畳み込み層を加える点ではDCGANと共通だが、G学習の過程で生成ネットワークと識別ネットワークに層を追加しながら、対応する画像の解像度を上げていくという特徴がある。レイアウトや配色は複雑な情報は必要とせず、解像度の低いレベルで重要な情報を多く含んでいると考えられるので、それらより捉えた学習ができる手法であるPGGANを採用する。PGGANでは式(4)に示す損失関数を用いて識別器の学習を行う。

$$L = \underbrace{\mathbb{E}_{\hat{x} \sim \mathbb{P}_g} [D(\hat{x})] - \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)]}_{\text{critic loss}} + \lambda \underbrace{\mathbb{E}_{\hat{x} \sim \mathbb{P}_g} [\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1]^2}_{\text{Gradient penalty}} \quad (3)$$

ここで、 $D(x)$ は画像 $x$ に対する識別器の予測結果であり、 $\mathbb{P}_g$ は生成器によって生成された画像集合、 $\mathbb{P}_r$ は学習データとなる画像集合である。本手法での評価値算出は、学習済みの識別器に対して、(1)式で表すcritic lossを用いる。実際にデザイン機構によって生成される画像 $g$ を識別器に入力したときの出力の値と、学習に用いたバナー画像を変換した重要度マップの識別器の出力の平均のユークリッド距離をそれぞれの画像の評価値 $s$ として用いる(式4)。

$$s = \left| [D(g)] - \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)] \right| \quad (4)$$

### 4.1 レイアウト評価器

レイアウト評価器は入力として与えられる文字やビジュアルを含む各素材の大きさと配置に関してバナー広告画像としての妥当性という観点から評価を行う。評価機構を構築するため、入力として与えられるグラフィックデザインの画像から視覚的に重要度を示した重要度マップに変換する(図3)。マップ画像の変換に事前学習済みのVisual Importance Model[11]を用いる。これは人間がグラフィックデザインを見たときの注目領域を深層学習を用いて構築したもので、同様のものとして前述の顕著性マップ[12]があるが、これは主に自然画像を対象としていて、事前知識を含まないボトムアップ効果によるモデルであるため、顕著性マップよりも、Visual Importance Modelの方がより適切にバナー広告画像のレイアウトの特徴を表していると考えられる。

学習データとするバナー広告画像の視覚的重要度を表すマップ画像を学習データとして、画像を生成するようにGANを学習させる。そこで得られる識別器は素材のレイアウトに関してバナー広告画像としての妥当性という観点から評価できるようになると考えられる。評価機構のためのGANの機構を図4に示す。



図 3 Visual Importance Model による重要度マップへの変換

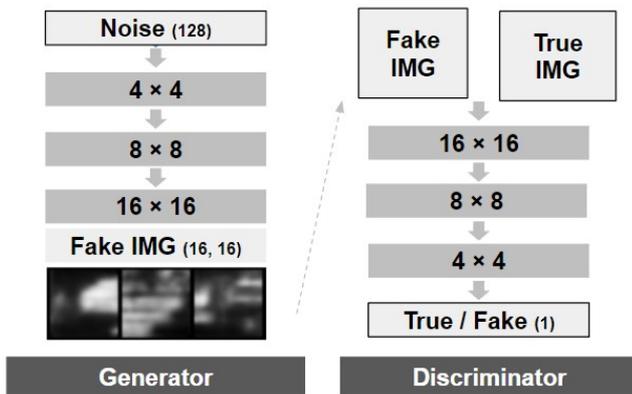


図 4 レイアウト評価器構築のための GAN

## 4.2 配色評価器

配色評価器は、入力として与えられる画像に対して、全体の配色の適切さに関して評価を行う。配色評価器は深層ニューラルネットワークモデルとして実現し、GAN の識別器として学習を行う。ここでは、既存のバナー画像を  $16 \times 16$  画素の低解像度にダウンサンプリングした画像を学習データとする(図 5)。ダウンサンプリングするのは、画像の局所的な色の特徴を無視して、バナー広告画像の全体的な配色を評価するためである。ダウンサンプリングには求めたい画素値を元画像の周辺  $4 \times 4$  画素の輝度値を用いて計算を行うバイキュービック補間手法を用いる。生成器は低解像度の画像を生成するよう学習させることによって、そこで得られる識別器は全体の配色に関して広告としての適切さを評価できるようになる。ここで学習された識別器を配色評価器として用いる。評価機構のための GAN の機構を図 4.6 に示す。



図 5 バイキュービック補間によるダウンサンプリング

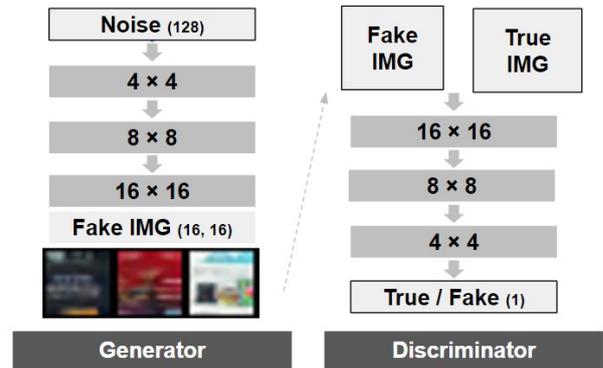


図 6 配色評価器のための GAN

## 5 評価実験

4 章で述べた DCGAN と PGGAN を Web 上に存在する解像度  $300 \times 250$  画素のバナー画像 7200 枚を学習データとしてバッチサイズ 4 でそれぞれレイアウト評価器と配色評価器の学習を行った。バナー画像の収集には Moat.com を利用した。学習に使用したコンピュータの性能は以下の通りである。また、学習の際の最適化関数は Adam [13] を用いた。

- CPU: Intel(R) Xeon(R) Silver 4110
- GPU: NVIDIA Geforce RTX 2080 Ti
- Memory: 94GB
- Deep Learning Framework: PyTorch

### 5.1 レイアウト評価器の実験

以下の 3 モデルについて精度比較を行う。

- LVIPG: Visual Importance Model によって変換し PGGAN によって評価器の構築を行ったレイアウト評価器
- LSLPG: 顕著性マップによって変換し PGGAN によって評価器の構築を行ったレイアウト評価器
- LVIDC: Visual Importance Model によって変換し DCGAN によって評価器の構築を行ったレイアウト評価器

以下にその実験手順を示す。表 1 に示す素材を入力として、3.1 に示すレイアウトに関するパラメータ  $p^c$  をすべて乱数によってそれぞれが重複を許さない形で決定してバナー画像を 10000 枚生成する。これらを学習済みの LVIPF に入力し、レイアウト評価器の出力値高い順に並べ替えたときの、1, 201, 401, ...9801 番目に対応する画像と最下位である 10000 番目の画像の合計 51 枚の画像を抽出する。それらの画像を対象に「バナー広告のレイアウトとして適切だと思いますか?」という質問に対して 5 段階評価を行ってもらい、被験者の主観評価の値と LVIPG, LSLPG, LVIDC のそれぞれのレイアウト評価値との間に相関があるかを調べた。

表 1 評価実験に用いるレイアウト素材画像

Image	
Logo	<b>I-CORP</b>
Ad Text	<b>濁り仕立て緑茶</b>
button	Buy Now

## 5.2 レイアウト評価器の結果

各モデルの評価値の高い画像上位 5 枚を以下の図 7-図 9 に示す。また、被験者の評価値の平均値を正規化した値と各評価器のモデルの評価値を正規化した値の散布図を図 10 に示す。各モデル主観評価値と各モデルのレイアウトの評価値との関係についてピアソンの相関係数を用いて比較したところ LVIPF が一番精度が高かった。このことから、レイアウトの評価器は Visual Importance Model によって変換し、PGGAN によって評価器を構築したモデルが一番適していることが示された。

表 2 主観評価とレイアウト評価器の各モデルとの相関係数

Model	ピアソンの相関係数
LVIPG	<b>0.59</b>
LSLPG	0.18
LUIDC	0.19



図 7 LVIPG の評価値の上位 5 枚とその順位



図 8 LSLPG の評価値の上位 5 枚とその順位



図 9 LUIDC の評価値の上位 5 枚とその順位

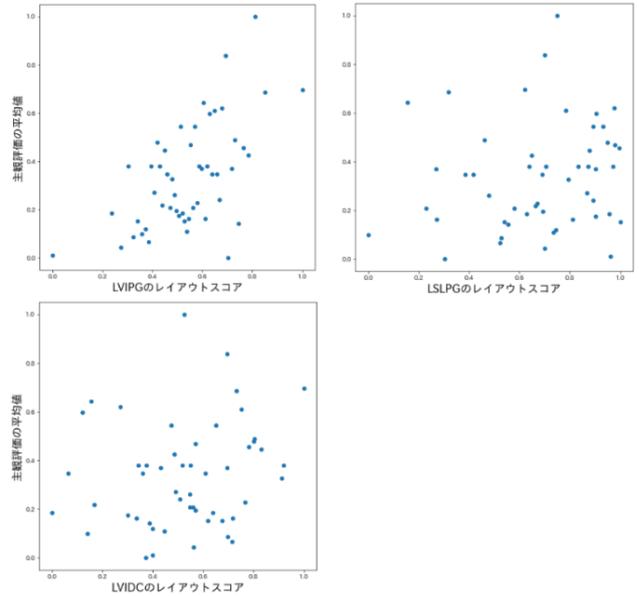


図 10 被験者の評価値の平均値と各レイアウト評価器の評価値の散布図

## 5.3 配色評価器の実験

以下の 2 つのモデルの精度の比較検討を行った。

- CPG: バイキュービック補間によって変換し、PGGAN を用いて構築した配色評価器
- CDC: バイキュービック補間によって変換し、DCGAN を用いて構築した配色評価器

以下にその実験手順を示す。配色評価器においても同様に、表 1 に示す素材を入力として、3.1 に示すレイアウトに関するパラメータ  $p_{color}$  をすべて乱数によってそれぞれが重複を許さない形で決定してバナー画像を 10000 枚生成する。実験 5.1 でレイアウト評価器の出力の値一番高かったレイアウトを利用して、これらを学習済みの配色評価器に入力し、レイアウト評価器の出力値が高い順に並び変えたときの、1, 201, 401, ...9801 番目に対応する画像と最下位である 9999 番目の画像の合計 51 枚の画像の画像を抽出する。これらの画像を対象に「バナー広告の配色として適切だと思いますか?」という質問に対して 5 段階評価を行ってもらい、被験者の主観評価の値と CPG, CDC のそれぞれの配色評価値との間に相関があるかを調べる。

## 5.4 配色評価器の結果

各モデルの評価値の高い画像上位 5 枚を以下の図 11, 図 12 に示す。また、被験者の評価値の平均値を正規化した値と各評価器のモデルの評価値を正規化した値の散布図を図 13 に示す。主観評価値と各モデルの配色の評価値との関係についてピアソンの相関係数を用いて比較したところ CDC が一番精度が高かった。このことから、配色の評価器はバイキュービック補間によって変換し、DCGAN を用いて構築したモデルが一番適していることが示された。

表 3 主観評価と配色評価器の各モデルとの相関係数

Model	ピアソンの相関係数
CPG	0.58
CDC	<b>0.73</b>



図 11 CPG の評価値の上位 5 枚とその順位



図 12 CDC の評価値の上位 5 枚とその順位

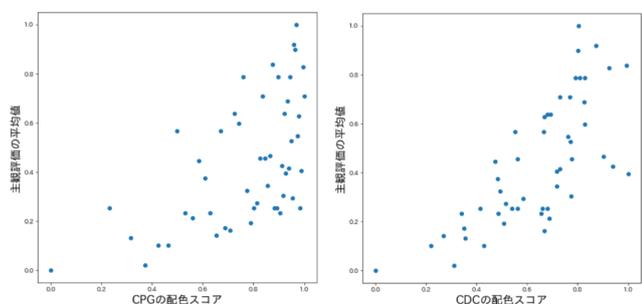


図 13 被験者の評価値の平均値と各配色評価器の評価値の散布図

### 5.5 総合評価の実験手順

レイアウトの評価に対しては Visual Importance Model によって変換し、PGGAN によって評価器を構築したモデルが一番精度が高く、配色の評価に対してはバイキュービック補間によって変換し、DCGAN を用いて構築したモデルが一番精度が高いことが示されたが、それぞれを組み合わせるときの総合的な印象が一番高いモデルの組み合わせを調査した。比較対象としては以下のモデルの組み合わせである。

- LVIPG + CPG: Visual Importance Model によって変換し、PGGAN によって評価器の構築を行ったレイアウト評価器と PGGAN を用いて構築した配色評価器の組み合わせ
- LSLPG + CPG: 顕著性マップによって変換し、PGGAN によって評価器の構築を行ったレイアウト評価器と PGGAN を用いて構築した配色評価器の組み合わせ
- LVIDC + CPG: Visual Importance Model によって変換し、DCGAN によって評価器の構築を行ったレイアウト評価器と PGGAN を用いて構築した配色評価器の組み合わせ
- LVIPG + CDC: Visual Importance Model によって変換し、PGGAN によって評価器の構築を行ったレイアウト評価器と DCGAN を用いて構築した配色評価器の組み合わせ

以下にその実験手順を示す。5.1 のにおいて 1, 2001, 4001, 6001, 8001, 10000 番目に対応する 6 組の  $p^l$  を利用して、それぞれのレイアウトに対して、 $p^c$  を乱数によって設定した画像を 10000 枚ずつ生成し、それらを CPG の配色評価器に入力し、それぞれのレイアウトごとに配色評価器の評価値が高い順に並び変えたときの、1, 2001, 4001, 6001, 8001, 10000 番目の画像の合計 36 枚抽出する。被験者にそれらの画像をランダムに掲示し、「総合的にバナー広告としてふさわしいと思いますか？」という質問に対して 5 段階評価を行ってもらい、それぞれのモデルの組み合わせによって算出されるグラフィックスコアに対して被験者の主観評価の値と提案モデルの出力との間に相関があるかを調べる。

### 5.6 総合評価器の実験結果

各モデルの評価値の高い画像上位 5 枚を以下の図 14-図 17 に示す。また、被験者の評価値の平均値を正規化した値と各評価器のモデルの評価値を正規化した値の散布図を図 19 に示す。主観評価値と各モデルの評価値との関係についてピアソンの相関係数を用いて比較した。LVIPG + CDC の相関係数が一番大きくなり、総合の評価においても各観点からの評価と同様に、レイアウトの評価器には Visual Importance Model によって変換し、PGGAN によって評価器の構築を行ったモデルと配色の評価器にはバイキュービック補間によって変換し、DCGAN を用いて構築した評価器を用いた組み合わせの精度が高かった。

表 4 手順 4 における評価器のモデルの組み合わせ

	レイアウトの評価器	配色の評価器	ピアソンの相関係数
モデル 1	LVIPG	CPG	0.73
モデル 2	LSIPG	CPG	0.43
モデル 3	LVIDC	CPG	0.30
モデル 4	LVIPG	CDC	<b>0.79</b>



図 14 LVIPG+CPG の評価値の上位 5 枚とその順位



図 15 LSLPG+CPG の評価値の上位 5 枚とその順位



図 16 LVIDC+CPG の評価値の上位 5 枚とその順位



図 17 LVIDC+CDC の評価値の上位 5 枚とその順位

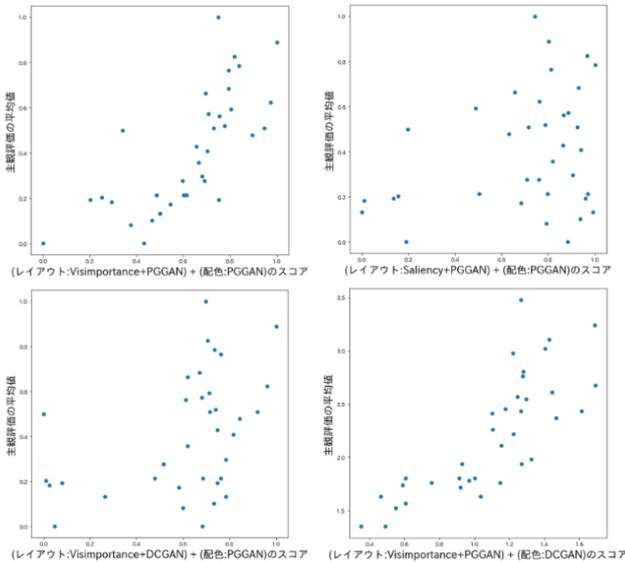


図 18 被験者の評価値の平均値と各モデルの評価値の散布図

## 6 ま と め

本論文は、グラフィックデザインとしての妥当性を「レイアウト」と「配色」という観点から評価を行うモデルを構築し、グラフィックデザインを自動生成する手法を開発した。

グラフィックデザインとしての妥当性の評価を行う評価機構はレイアウトの評価器には Visual Importance Model によって変換し、PGGAN によって評価器の構築を行ったモデル。配色の評価器にはバイキュービック補間によって変換し、DCGAN を用いて構築した評価器を用いたときが各観点からの評価、総合的な評価のどちらにおいても一番精度が高く、グラフィックデザイン生成システムに一番適していることが示された。

本論文においてはパナー広告を対象としたが、対象物は基本的に他のグラフィックデザインに適用可能と考えられる。そのため、活用方法としては、グラフィックデザインのデザインスキルや経験があまりない人がグラフィックデザインを作成する際にシステムがユーザに候補群をいくつか提示し、その中の一つをユーザがよりよく仕上げていくようなデザインツールとしての活用が考えられる。また、グラフィックデザインとしての妥当性を評価することが出来る評価機構を開発したことにより、評価機構の評価値を活用し、既存のデザインをよりよくするために「どの要素をどれくらい動かした良いか」、「どの素材の色をどれくらい変えた方が良いか」ということをユーザに伝える最適化ツールとしても活用できると考えられる。

## 謝 辞

本研究は JSPS 科研費 19H04219 の助成を受けたものです。

## 文 献

- [1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.
- [2] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10684–10695, 2022.
- [3] Jianan Li, Jimei Yang, Aaron Hertzmann, Jianming Zhang, and Tingfa Xu. Layoutgan: Generating graphic layouts with wireframe discriminators. *arXiv preprint arXiv:1901.06767*, 2019.
- [4] Xinru Zheng, Xiaotian Qiao, Ying Cao, and Rynson WH Lau. Content-aware generative modeling of graphic design layouts. *ACM Transactions on Graphics (TOG)*, Vol. 38, No. 4, pp. 1–15, 2019.
- [5] Sreekanth Vempati, Korah T Malayil, V Sruthi, and R Sandeep. Enabling hyper-personalisation: Automated ad creative generation and ranking for fashion e-commerce. In *Fashion Recommender Systems*, pp. 25–48. Springer, 2020.
- [6] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [7] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, Vol. 26, , 2013.
- [8] 伊藤博臣. エンジニアのための理論でわかるデザイン入門 (Think IT BOOKS). インプレス R&D/インプレスビジネスメディア, 2017.
- [9] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.
- [10] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.
- [11] Zoya Bylinskii, Nam Wook Kim, Peter O'Donovan, Sami Alsheikh, Spandan Madan, Hanspeter Pfister, Frédo Durand, Bryan C. Russell, and Aaron Hertzmann. Learning visual importance for graphic designs and data visualizations. *CoRR*, Vol. abs/1708.02660, , 2017.
- [12] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 20, No. 11, pp. 1254–1259, 1998.
- [13] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.