

## 学位論文及び審査結果の要旨

横浜国立大学

氏名	崎濱 栄治
学位の種類	博士（経営学）
学位記番号	国府博甲第65号
学位授与年月日	令和4年3月24日
学位授与の根拠	学位規則（昭和28年4月1日文部省令第9号）第4条第1項及び横浜国立大学学位規則第5条第1項
研究科(学府)・専攻名	国際社会科学府経営学専攻
学位論文題目	コンピュータービジョンと機械学習の統合によるモバイル広告効果の要因分析手法 A Method for Analyzing the Contribution of Mobile Advertising Effectiveness through the Integration of Computer Vision and Machine Learning
論文審査委員	主査 横浜国立大学 鶴見 裕之 教授 横浜国立大学 本橋 永至 准教授 横浜国立大学 寺本 高 教授 横浜国立大学 田名部 元成 教授 横浜国立大学 井上 徹 教授

### 論文の要旨

日本における2020年の総広告費は新型コロナウイルス感染症の影響を受け前年度88.8%と落ち込んだが、インターネット広告は前年度105.9%と、新聞やテレビメディア広告が減少傾向にある中で唯一増加に傾向にあり[電通 20]、これは日本だけではなく海外も同様である[電通 21]。インターネット広告は、大別するとディスプレイ広告と検索連動型広告の二つが存在している。

Webサイト上や動画再生時、アプリ実行時に表示される広告はディスプレイ広告と呼ばれ、その多くがバナーと呼ばれる画像形式であるためバナー広告とも呼ばれる。Webサイトやアプリに掲載される場合は画像、動画といった形式や、コンテンツ間に配置するインフィード広告といった表示形式等、様々な種類が存在する。また動画サイトの場合は、動画再生前又は途中で表示される広告もディスプレイ広告の一種である。インターネット広告の初期からある広告形式であり、現在も広く使用されているが、近年は検索連動型広告のシェアがディスプレイ広告を上回っている。ディスプレイ広告は基本的には従来の屋外広告やテレビ広告と近い利用者のアクセスを待つ受動型の性質を有することから、ニーズが固まっていない潜在層の利用者向けに使用されることが多い。

一方で検索連動型広告とは、利用者が検索エンジンに入力したキーワードに連動した広告が表示される形式である。ディスプレイ広告がWebサイトに表示される一方、検索連動型広告は検索結果画面の上部に表示される。このようにニーズが明確な利用者を対象とする形式のため、ニーズの顕在層向けに使用されることが多く、有効性が高いとされている。なお、検索連動型広告をリスティング広告と言い換えることもあれば、ディスプレイ広告と検索連動型広告の双方を含めてリステ

イング広告と呼ぶこともある。

インターネット広告は、これら検索連動型広告やバナー広告・ディスプレイ広告と呼ばれる画像広告が一般的であったが、通信環境の改善、高性能なモバイル端末が広く普及したことを背景に、動画広告も活用されるようになってきている。[電通 20] によると、2020 年の動画広告の市場規模は 3,862 億円(前年比 21.3% 増) であり、検索連動型広告(前年比 1.6%増)とディスプレイ広告(前年比 3.4%増)と比較して大幅に成長している。動画広告は、1. 情報やイメージが印象に残りやすい、2. テキスト広告よりも多くの情報を伝えられる、3. 拡散力があり話題になりやすい、といった特徴を備えている。

本研究では、インターネット広告における画像広告・動画広告を題材とし、テクノロジーを援用することで人の創造力を更に発揮できる可能性を提示し、現時点で実装可能な「人と機械との協業」方法について示す。

検索連動型広告(リスティング広告)を対象とした事例として、広告効果が高い(クリックされやすい、商品購入に繋がりやすい等)とされるキーワードを特定することで、人間がより魅力的な広告文を効率よく作成できる支援方法について Google 社による事例が公開されている [GoogleJapan 18]。

コンピュータービジョンを動画素材に適用し特徴量と教師データとの関係を明らかにした研究として [Kaminski 20] がある。同研究では、クラウドファンディングサイトを対象とし、説明文のテキスト、プレゼンテーション動画の特徴量から、クラウドファンディングの成否を複数の機械学習手法を用いて予測するモデルの提案を行った。動画の音声情報は Google Speech-to-Text API を用いてテキスト情報に変換し、動画は VIA によるタグ付けを行っている。また、特徴量の重要度は罰則付きのロジスティック回帰の結果から得られるが、交互作用については今後の課題とされている。

インターネット広告研究では、広告効果の予測モデル構築は予測精度の向上に主眼が置かれ、近年では深層学習を活用してテキスト情報や画像や動画を包括して特徴量とする研究が進展している [Xia 20]。また、画像広告や動画広告の素材そのものを自動生成する研究も進んでいる [谷口 17]。しかしながら、モバイル広告のデータを対象とした「人と機械の協業」による効果的な広告クリエイティブ制作の支援を目的とする研究は、多く見当たらず研究のホワイトスペースが存在すると考えた。

具体的な研究内容として、大手 IT 企業が提供するコンピュータービジョンサービス、機械学習の一つであるアンサンブル学習、ブラックボックスとも言われる機械学習モデルの解釈を試みる説明可能 AI(Explainable AI: XAI)、自然言語処理を組み合わせることで、各特徴量の重要度だけでなく、交互作用も含めルールベースでアンサンブル学習モデルの解釈を可能とする。また、本研究の枠組みは、広告効果に相当する教師データ(売上金額、販売数量、人による評価集計等) が得られるタスクであれば、映像・画像・テキスト・音声を組み合わせて取り扱うことができることから、販促用ポスター広告、駅やオフィスビルにおけるデジタルサイネージ広告、タクシー内の動画広告など広告他にも、TV コマーシャル、ドラマ・映画の予告編といった幅広い映像コンテンツの評価にも適用することが可能であると考えた。

近年、深層学習関連の研究成果が汎用的なライブラリとして組み込まれ、画像や動画について分

析することが比較的容易になったことを背景に、深層学習研究や画像認識技術が発展している。事例としては、[Schroff 15]が画像認識で人間の精度を超える研究結果を報告したことが挙げられる。こうした背景の元、本研究では初めに、初めに画像広告や動画広告の素材内に存在している物体の名称、色彩、動画であれば場面転換の回数といった情報を人が解釈可能な形式で抽出する。また、画像広告と動画広告の素材から情報を抽出する際には、コンピュータービジョンを利用する。

本研究では、第3章において Microsoft Computer Vision API による画像広告の分析を、第4章において Google Cloud Vision による動画広告の分析を行った。Amazon, Google, Microsoft といった外資大手 IT 企業は保有するビッグデータを活かして、画像・動画の分析サービスを展開している。分析結果から人間が解釈できるキーワードや数値として、複数の特徴量が提供される。人間であれば性別や年齢、支配的な色調、画像の大きさといった分析結果を瞬時に確認することができる。これらサービスを利用するメリットは、自ら画像や動画などの教師データを大量に用意し、学習モデルを構築する必要がないことがあげられる。更に、第4章では抽出した文字情報に対して潜在ディリクレ分布法(Latent Dirichlet Allocation:LDA)によるトピック分布の推定を行い、同分布についても特徴量として考慮している。

次に、広告効果(広告クリックやアプリのダウンロード、商品購入)に対して重要度が高い特徴量を特定する必要がある。このために、第3章では教師データとして広告効果に広告クリックを第4章ではコンバージョン(広告アプリのダウンロード)を利用し、特徴量としてコンピュータービジョンによる素材の抽出結果と広告配信に関するログデータによるアンサンブル学習モデルの構築を行った。アンサンブル学習は複数の(弱)学習器からなる機械学習手法である。本研究では、構築したモデルを人が解釈することで、画像広告や動画広告の広告素材を制作する際に効率的な示唆を得る「人と機械の協業」を目指している。よって、説明可能 AI(Explainable AI:XAI)分野の成果から、可読性が高いアソシエーション・ルールによる解釈が適用可能とされるアンサンブル学習手法を用いた。第3章では勾配ブースティング決定木(Gradient Boosting Decision Tree:GBDT) [Friedman 01,02]を、第4章では XgBoost(eXtreme Gradient Boosting trees) [Chen 16a, Zhang 18]を用いた。また、説明可能 AI の手法としては、第3章は inTrees(Interpreting Tree Ensembles with inTrees)[Deng 19]を第4章は defragTrees[Hara 18]を利用した。

結果について、第3章では画像広告を、第4章では動画広告を対象としてアンサンブル学習と機械学習の解釈手法から交互作用を含めた効果検証の方法を提案した。広告配信のログに関しては情報に差はなく、それぞれの広告素材に応じた特徴量の検討を行った。なお、インターネット広告において広告が配信される広告掲載場所は広告効果と密接に関連している。例えば、Web サイトの最上段であれば表示されて人の目に触れる可能性が高い一方、Web サイトの最下段ではスクロールを続けなければそもそも表示されないため、広告の素材がどれだけ優れていてもその効果は限定的となる。このため、本研究においては対象データを特定の掲載場所とすることで広告素材の違いによる広告効果を測定した。画像広告と動画広告に関して、コンピュータービジョンによる特徴量抽出の結果を比較すると、画像広告で利用できる特徴量の数は動画広告より格段に少ない。そのため、第3章では色彩に関する特徴量を利用した。色彩に関する重要度は、28.05%と tag(画像内のオブジェクト)、キャンペーン(広告対象)に続いて3番目となった。一方、動画広告については含まれる人、動物、物のオブジェクトに加えて、オブジェクトの位置情報も含めると取得可能な特徴

量が膨大である。このため、人が解釈してクリエイティブ制作に活用しやすい特徴量を選択した。具体的には、動画に含まれるオブジェクトのラベルの他に、場面転換を表すシーン数、動画広告の長さ、広告内で表示されたテキスト情報を利用した。

予測精度の改善が目的であれば利用できる特徴量は多ければ多いほど望ましく、[Chen 16b]のように、CTR 予測モデル構築の際に、広告配信ログに加えて広告画像のピクセル情報も Deep Learning の特徴量として利用する方法がある。説明可能 AI(Explainable: XAI)の技術を利用して、広告画像のどの部分が CTR 予測に対して貢献しているか視覚情報として提供することができる。ただし、同手法によるフィードバックをクリエイターが理解する際には、広告画像を一枚ずつ目視で確認して広告効果への貢献が大きい領域に何が描かれているか判断する必要がある。本研究では、広告内に何が表現されているか、人が判断する代わりにコンピュータービジョンによってキーワードとして取得することを提案した。外資 IT 大手を筆頭に画像や動画を対象としたコンピュータービジョンサービスが展開されており、ビジネスにおいて如何に活用していくかというフェーズに突入しているため、本研究の分析の枠組みについても実現性が高いものであるといえる。

機械学習手法の分野は研究が日々進んでおり、改善された機械学習、自然言語処理、コンピュータービジョン、説明可能 AI を用いることでより一層精度の高い結果が得られる可能性が高い。この点は、本研究の限界であるが、これら複数の手法を組み合わせるスキームは「人と機械の協業」において様々な分野に応用可能であると考えられる。

インターネット広告はマーケティング活動の手段の一つであり、広告をクリックするとアプリストアや商品詳細に関するランディングページ(LP:Landing Page)などに遷移し、アプリのダウンロードや商品購入、サービス申込みといったコンバージョンへ誘導することが目的とされる。本研究の応用として、これら広告クリック後の遷移先の Web サイトが画像広告、動画広告のクリエイティブとマッチしているかどうかについても示唆が得られる可能性がある。単純なコンバージョン数や Web サイトの滞在時間だけではなく、Web サイトのクリエイティブについてコンピュータービジョンによる物体検出や自然言語処理によるテキスト分析を行い特徴量として利用することも考えられよう。

[Chen 16a] Chen,T. and Guestrin,C.: XgBoost:A scalable tree boosting system, Proceedings of 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.785–794 (2016)

[Chen 16b] Chen,J.,Sun,B.,Li,H.,Lu,H.and Hua,X.:Deep CTR Prediction in Display Advertising,Proceedings of the 24th ACM international conference on Multimedia,pp.811-820(2016)

[Deng 19] Deng,H.:Interpreting tree ensembles with inTrees,International Journal of Data Science and Analytics,Vol.7,pp.277-287(2019)

[Friedman 01] Friedman,J.H.:Greedy function approximation:A gradient boosting machine, Annals of Statistics,Vol.29,pp.1189–1232(2001)

[Friedman 02] Friedman,J.H.:Stochastic gradient boosting,Computational Statistics & Data Analysis,Vol.38,No.4,pp.367–378(2002)

[GoogleJapan 18] Google Japan, 機械学習のその先に初めの一步を踏み出そう (2018):[https://www.youtube.com/watch?v=jN\\_pREOP1GU](https://www.youtube.com/watch?v=jN_pREOP1GU)(access 2021/11/8)

[Hara 18] Hara,S.and Hayashi,K.:Making tree ensembles interpretable:A Bayesian model selection approach,Proceedings of the 21st International Conference on Artificial Intelligence and Statistics (AISTATS'18),pp.77-85(2018)

[Kaminski 20] Kaminski,J.C. and Hopp,C.:Predicting outcomes in crowdfunding campaigns with textual,visual and linguistic signals, Small Business Economics,Vol.55,pp.627-649(2020)

[Schroff 15] Schroff,F.,Kalenichenko,D.and Philbin,J.:FaceNet:A Unified Embedding for Face Recognition and Clustering,arXive preprint,1503.03832v3(2015)

[Xia 20] Xia,B.,Seshime,H.,Wang,X.,and Yamasaki,T.:Click-through rate prediction of online banners featuring multimodal analysis,International Journal of Semantic Computing, Vol.14, No.1,pp.71-91(2020)

[電通 20] 株式会社電通グループ,2020 年日本の広告費(2021)

[電通 21] 株式会社電通グループ:電通グループ「世界の広告費成長率予測 (2020~2022)」改定版を發表 <https://www.group.dentsu.com/jp/news/release/000509.html>(accessed 2021/11/19)

[谷口 17] 谷口和輝,大田和寛,山口光太:人工知能による新しい広告クリエイティブ,人工知能学会誌,Vol.32,No.4,pp.501-508(2017)

## 審査結果の要旨

企業のマーケティング・コミュニケーションにおいて、インターネット広告の重要性は年々高まっており、特にスマートフォン等に表示されるモバイル広告や動画広告は、通信環境の向上や高性能なモバイル端末の普及により、急速に拡大している。一方、画像解析や機械学習のテクノロジーの進歩は目覚ましく、それらがインターネット広告の効果の向上に貢献している。本論文の目的は、インターネット広告における画像広告・動画広告を題材として、画像解析と機械学習のテクノロジーを援用することにより、広告クリエイティブ制作支援における「人と機械との協業」の方法の可能性を示すことである。

本論文は以下の 5 章から構成される。

第 1 章 はじめに

第 2 章 先行研究のレビュー

第 3 章 コンピュータービジョンによる画像広告要素のクリック訴求効果の検証

第 4 章 アンサンブル学習と LDA の統合による動画広告の要因分析

第 5 章 おわりに

第 1 章では、本研究の背景を整理し、本論文の目的と構成を示す。第 2 章では、本研究に関

連する先行研究がレビューされる。第3章では、モバイル広告を対象とした CTR 予測問題の枠組みで、広告画像における構成要素の貢献度を測定する手法を提案する。第4章では、動画広告から解釈可能なキーワードを抽出し、広告効果（コンバージョン）に対して、相互作用を考慮した上で重要な要素を特定する手法を提案する。第5章では、本研究の貢献と限界が整理される。

本研究では、画像広告と動画広告を対象として、アンサンブル学習と機械学習の解釈手法を用いて、相互作用を含めた効果検証の方法を提案した。本研究は、最先端のテクノロジーを活用したインターネット広告に関する研究として、学術と実務の双方において、貢献度の高い研究といえる。

本論文の一部は、『応用統計学』第48巻3号に「コンピュータービジョンによる広告画像要素のクリック訴求効果の検証」として、『人工知能学会論文誌』第36巻3号に「アンサンブル学習と LDA の統合による動画広告効果の要因分析」として、いずれも査読付き論文として掲載されている。また、本論文の一部の内容については、日本マーケティング・サイエンス学会研究大会にて学会報告を行っている。これらの実績は、本論文の質の高さを保証するものと捉えることができる。

以上のことから、本論文審査委員一同は、本学府の博士号審査基準①に照らして、崎濱栄治氏の学位請求論文「コンピュータービジョンと機械学習の統合によるモバイル広告効果の要因分析手法」が博士（経営学）の学位を授与するに値するものとして、判断する。

令和4年1月31日

審査委員主査	横浜国立大学大学院国際社会科学研究院教授	鶴見裕之
審査委員	横浜国立大学大学院国際社会科学研究院准教授	本橋永至
審査委員	横浜国立大学大学院国際社会科学研究院教授	寺本高
審査委員	横浜国立大学大学院国際社会科学研究院教授	田名部元成
審査委員	横浜国立大学大学院国際社会科学研究院教授	井上徹

参考；崎濱栄治氏の指導委員会の構成員は以下の通りである。

責任指導教員	横浜国立大学大学院国際社会科学研究院准教授	本橋永至
指導教員	横浜国立大学大学院国際社会科学研究院教授	鶴見裕之
指導教員	横浜国立大学大学院国際社会科学研究院教授	寺本高