

# 学術用語解説ウェブページにおける ページ構成要素を考慮した見易さの評定および理由提示

## Measuring and Explaining Visual Intelligibility of Web Pages on Academic Concepts considering Page Components

岡田心太朗<sup>†</sup> 曾田 耕生<sup>†</sup> 大賀 悠平<sup>†</sup> 宇津呂武仁<sup>††</sup> 河田 容英<sup>†††</sup>

<sup>†</sup> 筑波大学大学院 システム情報工学研究科・群 知能機能システム専攻・学位プログラム  
〒305-8573 茨城県つくば市天王台 1-1-1

<sup>††</sup> 筑波大学 システム情報系 知能機能工学域 〒305-8573 茨城県つくば市天王台 1-1-1  
理化学研究所 革新知能統合研究センター 〒103-0027 東京都中央区日本橋 1-4-1

<sup>†††</sup> (株) ログワークス 〒151-0053 東京都渋谷区代々木 1-3-15 天翔代々木ビル 6F

あらまし 本論文では、学術用語を解説するウェブページ群を対象として、深層学習を用いたページの見易さ自動評定、および評定理由の提示を行う。また、物体検出モデルを用いてページ構成要素の自動同定を行う手法を提案する。さらに、用語解説ページの画像データに加えて、ページ構成要素(タイトル・見出し、目次、解説文、図表、広告)の領域情報を考慮することにより、見易さ自動評定モデルの性能を改善するとともに、評定理由の具体化が行えることを示す。

キーワード 学術用語解説ウェブページ, 見易さ評定, 理由提示, ページ構成要素, ResNet, Grad-CAM, YOLO

## 1 はじめに

近年のモバイル端末の普及により、インターネットの存在はますます我々の暮らしに欠かせないものとなった。またその手軽さから、新しい事柄を学習する際にウェブコンテンツを利用することはもはや当たり前のこととなった。しかし、既存の検索エンジンを用いた学習コンテンツの検索においては、ページランキングの上位に表示されるウェブページが必ずしも学習者に適したコンテンツであるとは限らない。そのため学習者は、検索結果として挙げられた複数のウェブページを自ら比較し、学習に適したページを見つけ出す必要があり、この作業が快適な学習の妨げとなっている。そこで、本論文では、学術用語解説ウェブページを対象として、「学び易さ」に基づく順位付けを行うとともに、「学び易さ」の要因を特定することで、ウェブ学習の支援やウェブページの制作支援を行うことを目的とする。

本論文では、学術用語解説ウェブページの「学び易さ」に関する要因の一つとして「見易さ」に焦点を当て、深層学習モデルを利用した見易さの自動評定およびその理由提示を行う手順を示す。この際、特にページの画像情報に加えてページ構成要素(見出し、目次、解説文、図表、広告)の領域情報を考慮する手法を提案する。具体的には、まず参照用ウェブページデータセットとして、インターネット上の学術用語解説ウェブページを収集し、それらの見易さ評定およびページ構成要素の付与作業を手で行う。その後、作成した参照用データセットを元に、ResNet-50 モデル、および、Grad-CAM (Gradient-weighted

Class Activation Mapping) を用いて、ページ画像のみを考慮した見易さ自動評定および理由提示を行う。その結果、一定以上の性能で見易さ自動評定、および、理由提示を行うことができた。次に、汎用物体検出モデルを利用して、ページ構成要素の同定を自動で行う手法を提案する。具体的には、大規模データセットを用いて事前訓練された物体検出モデルである YOLOv5 モデルを利用することにより、概ね高い精度で自動同定を行うことができた。最後に、ページ構成要素の自動同定結果を利用して見易さ自動評定、および、その理由提示を行う。ページ構成要素を考慮した見易さ自動評定タスクでは、ResNet-50 の部分アーキテクチャを利用してモデルを設計し、ページ画像に加えてページ構成要素の領域情報を入力することで、自動評定精度を向上させることができた。また、ページ構成要素を考慮した理由提示タスクにおいては、見易さ自動評定モデルの勾配を利用することで、自動評定結果に最も寄与したページ構成要素を特定した。これにより、見易さ自動評定に加えてページ構成要素を特定した理由提示を実現することができた。

## 2 参照用データセットの作成

### 2.1 対象学術分野および用語

本論文では、「ページデザインの見易さ」、「文章の分かり易さ」及び「全体評定」の評定基準の類似性に基づき主に理数系学術分野を対象とし、特に、「線形代数」・「解析」・「統計」・「力学」・「電磁気」・「化学」・「生物」・「医学」・「プログラミング」・「IT」の 10 分野を対象分野としてウェブページの収集を行った。ま

表 1 学術分野ごとの対象学術用語一覧

分野	学術用語数	対象とする学術用語	正例数および負例数		
			正例	負例	合計
線形代数	15	クラメルの公式, クロネッカーのデルタ, ノルム, 等 (計 15 語)	34	94	128
解析	15	$\varepsilon$ - $\delta$ 論法, コーシーの積分定理, テーラ展開, 等 (計 15 語)	38	100	138
統計	15	F 分布, コーシー分布, パレート分析, ボアソン分布, 等 (計 15 語)	81	132	213
力学	15	パスカルの原理, 水平投射, ケプラーの法則, 万有引力, 等 (計 15 語)	70	89	159
電磁気	15	ガウスの法則, ホイートストンブリッジ, ローレンツ力, 等 (計 15 語)	91	94	185
化学	15	同素体, 極性分子, カルボン酸, 共有結合, 典型元素, 等 (計 15 語)	57	85	142
生物	15	発酵, RNA, 中枢神経, マクロファージ, 受容体, 等 (計 15 語)	68	134	202
医学	15	チアノーゼ, バイオフィルム, リンパ腫, 血清, 外用薬, 等 (計 15 語)	148	160	308
IT	15	API, プロトコル, ディレクトリ, 拡張子, クエリ, 等 (計 15 語)	104	153	257
プログラミング	15	標準入力, コマンドライン引数, コンパイル, SDK, 等 (計 15 語)	77	136	213
合計	105	—	632	1,313	1,945

た, 各分野に属する学術用語として, 主に高校・大学レベルの用語の中から表 1 に示す 15 語を選んだ。

## 2.2 「見易さ」評定の個別因子

本論文では, 用語解説ウェブページの見易さ評定を人手で行うに当たって, 大きく分けて, 絶対条件と優先付け項目の 2 つを設け [8], 絶対条件を全て満たしており, かつ, 複数の優先付け項目を満たすウェブページを「見易い」と判定した。

## 2.3 ページ集合の作成手順および作成結果

まず, 表 1 に示す各学術用語を検索クエリとしてウェブ検索を行い, 検索結果上位 30 件のウェブページを収集する。その後, 収集した計 4,500 件のウェブページのうち, 本論文で対象とする複数の学術分野の検索結果に, 横断的に出現するドメインのウェブページを全て除外する<sup>1</sup>。結果として残った 1,945 件のウェブページについて, 前節で述べた基準に基づき, それぞれ見易さ評定を人手で行う。その結果, 各分野ごとに表 1 に示す数の正例・負例を収集することができた。

## 2.4 学術用語解説ウェブページ内のページ構成要素

本論文では, 学術用語解説ページの主要なページ構成要素として, 見出し, 目次, 解説文, 図表, 広告の 5 つを用いる。各ページ構成要素の定義を表 2 に, 例を図 1 に示す。

## 2.5 ページ構成要素データセットの作成手順および作成結果

2.3 節において人手によって見易さ評定を行った 1,945 件のウェブページについて, 前節で述べた定義に基づきページ構成要素の人手付与作業を行った。ページ構成要素の有無と, ページ全体での「見易い・見易くない」の人手評定結果との間の相関を測定するために, 表 3(a) の  $n_{ij}$  ( $i, j = 1, 2$ ) を用いたイエーツの  $\chi^2$  検定により相関の有無を判定した。

$$\chi^2 = \frac{N \left( \max(0, |n_{11}n_{22} - n_{12}n_{21}| - N/2) \right)^2}{n_{1.}n_{.1}n_{.2}n_{2.}}$$

この検定統計量  $\chi^2$  より求めた P 値<sup>2</sup>が 0.01 未満 (有意水準 1% で有意差あり) であった場合に, ページ構成要素の有無とページ全体での人手評定結果との間に相関があると判定した。その結果を表 3(b)~ 表 3(f) に示す。各用語解説ウェブページにおける目次, 解説文, 広告の有無は, そのページの見易さ評定結果と有意な相関があることが確認できた。

## 3 見易さ自動評定の評定理由

### 3.1 Grad-CAM

本論文では, 学術用語解説ウェブページの評定理由の分析手法として, Grad-CAM (Gradient-weighted Class Activation Mapping) [7] を用いて評定理由の可視化を行った [4], [5]。この手法の利点として, 訓練済みの CNN モデルの構造を変えずにモデルの判断根拠の分析が行える点や, 背景色や空白部分等, 画像中の広範囲の領域を評定理由とする可視化ができる点が挙げられる。一方で欠点としては, 畳み込み最終層の特徴量マップを用いている都合上, 入力画像と同程度の解像度で可視化を行うことができない点やモデル構造を変更した場合には, 理由可視化システム側も修正する必要がある点が挙げられる。

### 3.2 評定理由の提示結果の評価

前節の評定理由の提示手法を用いた可視化画像に対して, 評定理由の自動提示箇所と, 人手により「見易い」もしくは「見易くない」として提示された箇所との間の重複度合いに応じて, 表 4 に示す基準に基づいて 4 段階で評価を行った。評定理由提示結果の評価事例としては, 2 節で収集した 1,945 ページの学術用語解説ウェブページ集合から無作為に選定した 500 ページを用いた。自動評定モデルの自動評定値 (0~1) を幅 0.1 の範囲に分割し, 各範囲ごとの評価結果の詳細を表 5 に示す。自動評定モデルの自動評定値の全体においては, 約 68% の理由提示結

1: 一般的に同一ドメインのウェブページ間ではページの構成やデザインが類似しており, 人手による見易さの評定結果にも大きな偏りがあることから, このようなページを見易さ自動評定モデルの訓練に用いると「ウェブページの見易さ」モデルの訓練の妨げになると考えられる。

2: Scipy([https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.chi2\\_contingency.html](https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.chi2_contingency.html)) を用いて算出した。

表2 ページ構成要素の定義

ページ構成要素	定義
見出し	ページタイトルや見出し・小見出し。サイト名は含まない。
目次	同ページ内の内容を示す。サイト内のコンテンツを示す目次は含まない。
解説文	学術用語を解説する文章領域。図表内の文章や内容と関係のない文章(サイトや著者の説明、広告内の文章等)は含まない。
図表	解説用の図表。解説用でないイメージ画像等は含まない。
広告	Google AdSense 等による広告や、サイト独自の商品・サービスの購入および加入を促す広告。画像付きの別ページへのリンクは含まない。ヘッダーやサイドバーに位置する広告も含む。

表3 ページ構成要素の有無と「見やすい・見易くない」との相関

(a) 相関分析のための 2x2 分割表				(b) 見出し				(c) 目次			
	あり	なし	合計		あり	なし	合計		あり	なし	合計
見やすい	$n_{11}$	$n_{12}$	$n_{1\cdot}$	見やすい	765	3	768	見やすい	207	561	768
見易くない	$n_{21}$	$n_{22}$	$n_{2\cdot}$	見易くない	1,160	17	1,177	見易くない	147	1,030	1,177
合計	$n_{\cdot 1}$	$n_{\cdot 2}$	$N$	合計	1,925	20	1,945	合計	354	1,591	1,945
$\chi^2$ 検定による 相関の有無 (有意水準 1%)				相関なし (P 値=0.043)				$\chi^2$ 検定による 相関の有無 (有意水準 1%)			
								相関あり (P 値 < 0.001)			

(d) 解説文				(e) 図表				(f) 広告			
	あり	なし	合計		あり	なし	合計		あり	なし	合計
見やすい	766	2	768	見やすい	404	364	768	見やすい	213	555	768
見易くない	1,146	31	1,177	見易くない	601	576	1,177	見易くない	210	967	1,177
合計	1,912	33	1,945	合計	1,005	940	1,945	合計	423	1,522	1,945
$\chi^2$ 検定による 相関の有無 (有意水準 1%)				$\chi^2$ 検定による 相関の有無 (有意水準 1%)				$\chi^2$ 検定による 相関の有無 (有意水準 1%)			
相関あり (P 値 < 0.001)				相関なし (P 値=0.536)				相関あり (P 値 < 0.001)			

果が「4」または「3」と評価され、また、1件を除く全ての理由提示結果において「2」以上の評価となった。特に、自動評定値が0.0~0.1、または、0.9~1.0の範囲の事例では、「4」または「3」の評価の割合が比較的高いことが読み取れる。

評定理由の可視化例<sup>3 4 5 6</sup>および各例に対する評価を図2に示す。見易さ自動評定モデルの判定結果において、各ウェブページ内で相対的に「見やすい」と判断された領域が青色に、相対的に「見易くない」と判断された領域が赤色に染められている。図2(a)は、評定理由の提示結果の評価において最高の「4」と評価されたウェブページとその理由提示画像である。このウェブページの手見易さ判定においては、用語解説用の図や大きい目の文字で強調された表題、色付きの目次等が「見やすい」とされ、文字の細かい文章領域が「見易くない」とされた。そして、それらの領域が、見易さ自動評定モデルにおいても概ね同様に判断されていることが読み取れる。一方で、図2(b)、および、図2(c)に示す用語解説ウェブページでは、評定理由の提示結果の手評価が「2」、あるいは、「3」となった。

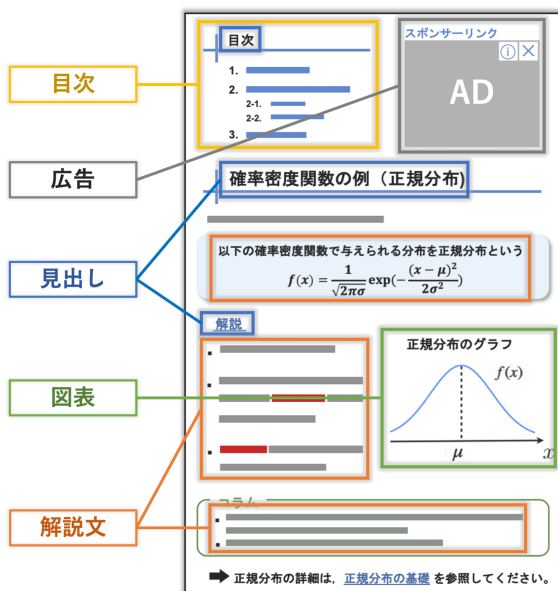


図1 ページ構成要素の例

3: 図2(a): [https://www.macromill.com/service/data\\_analysis/regression-analysis.html](https://www.macromill.com/service/data_analysis/regression-analysis.html)

4: 図2(b): <http://www2s.biglobe.ne.jp/~aaihara/circular.html>

5: 図2(c): <https://www.innopm.com/blog/2015/10/28/103108/>

6: 図2(d): [https://radiology.bayer.jp/training/livermri/useful\\_information/hepatoma/basics1/](https://radiology.bayer.jp/training/livermri/useful_information/hepatoma/basics1/)

これらの例では、「見易い」領域として目次や用語解説用の図、メニューバーなどが、また、「見易くない」領域として文字の細かい文章部分が、それぞれ挙げられるが、Grad-CAM による理由提示画像ではそれらの一部のみしか染められていないことが分かる。また、図 2(d) に示す用語解説ウェブページの理由提示画像では、人手により「見易い」もしくは「見易くない」として提示された箇所と、評定理由の自動提示箇所との重複領域はほとんど見られないため、「1」と評価した。

## 4 物体検出モデルによるページ構成要素の同定

### 4.1 YOLO[6]

本節では、物体検出モデルである YOLO [6] (YOLOv5<sup>7</sup>) を用いてページ構成要素の同定を自動で行う (図 3)。まず、入力画像を “Main Network” に入力し、サイズの異なる 3 つの出力テンソルを得る。これにより大きさの異なるあらゆる検出要素情報を抽出する。その後、これらのテンソルをそれぞれ異なる畳み込み層で処理することによって、検出した構成要素の種類、バウンディングボックスの位置、および、信頼度を得る。

### 4.2 訓練・評価手順および評価結果

本論文では、COCO 2017 データセット [2] を用いて事前訓練されたパラメータを初期値とする YOLOv5 モデルに対して、学術用語解説ページの訓練事例を用いた全パラメータの訓練を行った。訓練・開発・評価事例としては、表 6 に示す 5 種類の学術分野の組み合わせを用意し、5 分割交差検定によってモデルの訓練・評価を行った。また、モデルに入力するページ画像は、アスペクト比を固定した状態で横幅を 640px に縮小して使用した。モデルの評価時は、2 つのバウンディングボックスの重なり度合いを示す指標 IoU (Intersection over Union, 6.2 節参照) を使用し、参照用要素のバウンディングボックス、および、モデルが検出した要素のバウンディングボックスの IoU が 0.5 以上である場合を正解とした。

各ページ構成要素の参照用総数、および、ページ構成要素同定モデルの Average Precision (AP) を表 7 に示す。また、再現率-適合率曲線を図 4 に示す。ページ構成要素の種類間で AP の平均 (mean AP:mAP) は 0.75 であり、一定以上の性能によってページ構成要素が同定できた。特に、見出し、目次、図表の同定性能が相対的に高く、広告の同定性能が相対的に低くなった。ページ構成要素同定モデルが 0.5 以上の信頼度で出力した構成要素の例<sup>8,9</sup>を図 5 に示す。一部の構成要素では部分的な欠損があるものの、概ね全ての要素が適切に同定できている。

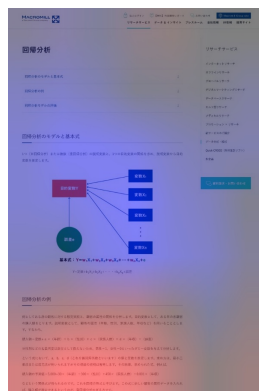
## 5 学術用語解説ウェブページの見易さ自動評定におけるページ構成要素の有効性

### 5.1 ニューラルネットワークのアーキテクチャ

本節では、ページの画像情報に加えてページ構成要素の領域

表 4 評定理由提示結果の評価基準

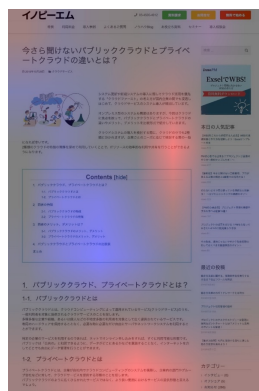
評価値	説明
4	人手による見易さ評定に寄与する領域と、Grad-CAM による評定理由の自動提示領域の 8 割以上が一致しており、かつ、「見易い/見易くない」の極性も一致している。(それ以外に、箇所は一致するが、極性は異なる領域があってもよい)
3	人手による見易さ評定に寄与する領域と、Grad-CAM による評定理由の自動提示領域の 6 割 ~8 割が一致しており、かつ、「見易い/見易くない」の極性も一致している。(それ以外に、箇所は一致するが、極性は異なる領域があってもよい)
2	人手による見易さ評定に寄与する領域と、Grad-CAM による評定理由の自動提示領域の 3 割 ~6 割が一致しており、かつ、「見易い/見易くない」の極性も一致している。(それ以外に、箇所は一致するが、極性は異なる領域があってもよい)
1	人手による見易さ評定に寄与する領域と、Grad-CAM による評定理由の自動提示領域が一致する割合が 3 割未満である。



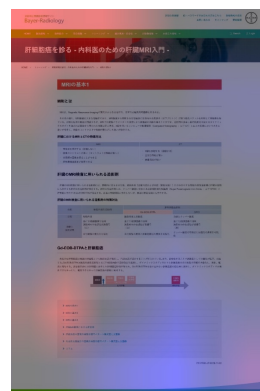
学術分野: 統計  
見易さの人手評定: 見易い  
見易さの自動評定確率 (0~1): 0.87  
(a) 評定理由提示結果の評価: 4



学術分野: 力学  
見易さの人手評定: 見易くない  
見易さの自動評定確率 (0~1): 0.09  
(b) 評定理由提示結果の評価: 3



学術分野: IT  
見易さの人手評定: 見易くない  
見易さの自動評定確率 (0~1): 0.48  
(c) 評定理由提示結果の評価: 2



学術分野: 医学  
見易さの人手評定: 見易い  
見易さの自動評定確率 (0~1): 0.19  
(d) 評定理由提示結果の評価: 1

図 2 Grad-CAM による見易さ評定理由の可視化画像に対する評価例

7: <https://github.com/ultralytics/yolov5>

8: 図 5(a): [http://otu4tokkun.net/kagakubuturi\\_1.html](http://otu4tokkun.net/kagakubuturi_1.html)

9: 図 5(b): <http://www.nteku.com/condenser/Default.aspx>

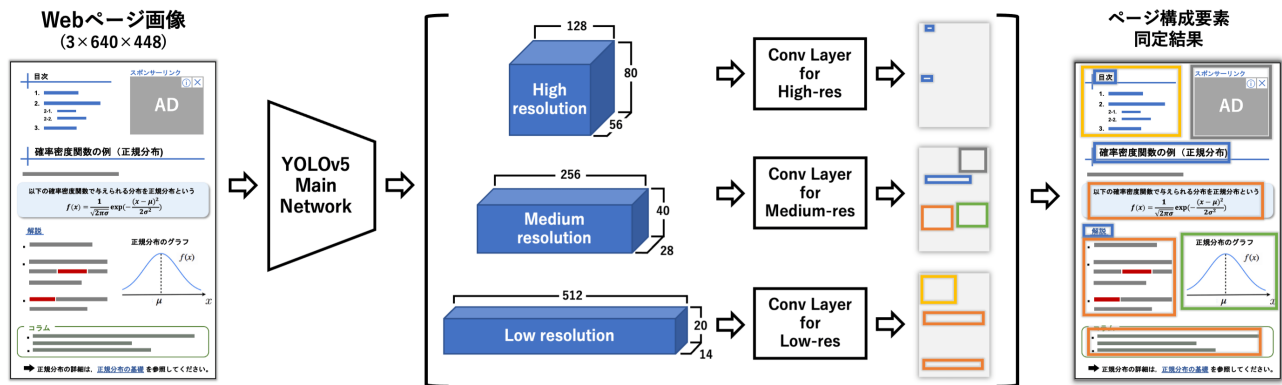


図3 YOLOv5 [6] を用いたページ構成要素同定の仕組み

表5 見易さ評定理由提示結果に対する人手評価結果

見易さ確率 (0 ~ 1)	人手評価値				合計 (%)
	4 (%)	3 (%)	2 (%)	1 (%)	
0.0 ~ 0.1	39 (21.7)	97 (53.9)	44 (24.4)	0 (0)	180 (100)
0.1 ~ 0.2	2 (4.8)	22 (52.4)	17 (40.5)	1 (2.4)	42 (100)
0.2 ~ 0.3	4 (9.5)	13 (31.0)	25 (59.5)	0 (0)	42 (100)
0.3 ~ 0.4	3 (15.0)	9 (45.0)	8 (40.0)	0 (0)	20 (100)
0.4 ~ 0.5	3 (12.5)	11 (45.8)	10 (41.7)	0 (0)	24 (100)
0.5 ~ 0.6	3 (12.0)	12 (48.0)	10 (40.0)	0 (0)	25 (100)
0.6 ~ 0.7	2 (9.1)	12 (54.5)	8 (36.4)	0 (0)	22 (100)
0.7 ~ 0.8	2 (9.5)	7 (33.3)	12 (57.1)	0 (0)	21 (100)
0.8 ~ 0.9	5 (15.2)	18 (54.5)	10 (30.3)	0 (0)	33 (100)
0.9 ~ 1.0	11 (12.1)	64 (70.3)	16 (17.6)	0 (0)	91 (100)
合計	74 (14.8)	265 (53.0)	160 (32.0)	1 (0.2)	500 (100)

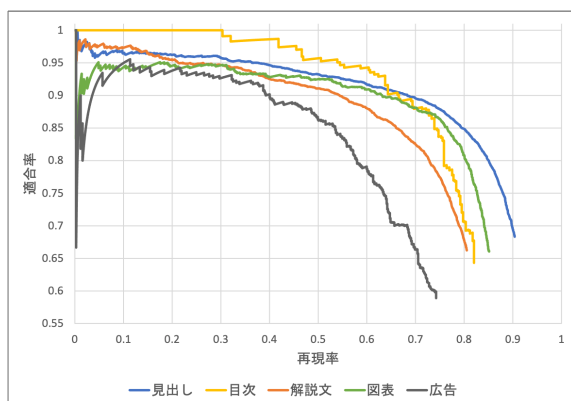


図4 ページ構成要素同定モデルの再現率-適合率曲線

情報も考慮した見易さ自動評定モデルの作成手順を示す。本論文で設計したアーキテクチャの概要を図6に示す。モデルには、ページ画像 (RGB) を表す3チャンネル分の入力、および、4節で作成した物体検出モデルにより検出されたページ構成要素の領域および信頼度情報を表す5チャンネル分の入力を与える。それらの3チャンネル分、および、5チャンネル分は、それぞれ別の畳み込みネットワークで処理する。これら2つのネットワーク基盤には、ImageNet データセットによって事前訓練済みの ResNet-50 モデル [1] を使用した。具体的には、ResNet-50 モデルの入力層から13個目の残差ブロックまでのアーキテクチャを複製し活用した。その後、2つのネットワークの出力を結合し、出力用の畳み込みネットワークに通すことで、ウェブページの見易さ確率を求める。出力用ネットワークでは、畳み込み層と平均プーリング層、全結合層をそれぞれ1層ずつ使用する。ベースラインモデルとしては、ResNet-50 モデルにおいてページ構成要素を考慮しないアーキテクチャ [8] を用いる。

## 5.2 訓練・評価手順

訓練・評価時に使用する用語解説ウェブページデータセットとしては、学術分野ごとに表6に示す5つの組み合わせを用意し、5分割交差検定によってモデルを評価した。前節で述べたアーキテクチャのパラメータの訓練を行う際、ページ画像 (RGB) を表す3チャンネル分のネットワークのパラメータとしては、ImageNet データセットによって事前訓練済みの ResNet-50 モデル [1] のパラメータをそのまま用い、パラメータの訓練は行わない。一方、ページ構成要素用の5チャンネル分のネットワークについては、ImageNet データセットによって事前訓練済みの ResNet-50 モデル [1] のパラメータを初期値としてパラメータの訓練を行う。また、出力用ネットワークのパラメータについては、無作為な初期値のもとで、パラメータの訓練を行う。また、モデルの入力となるウェブページ画像およびページ構成要素の領域を表すチャンネルは、ともに 224 × 224 のサイズにリサイズして使用する。

## 5.3 評価結果

前節の手順で訓練した見易さ自動評定モデルの再現率-適合率曲線を図7に示す。「見易い」事例、および、「見易くない」事例の両方において、ページ構成要素を考慮した見易さ自動評定モ

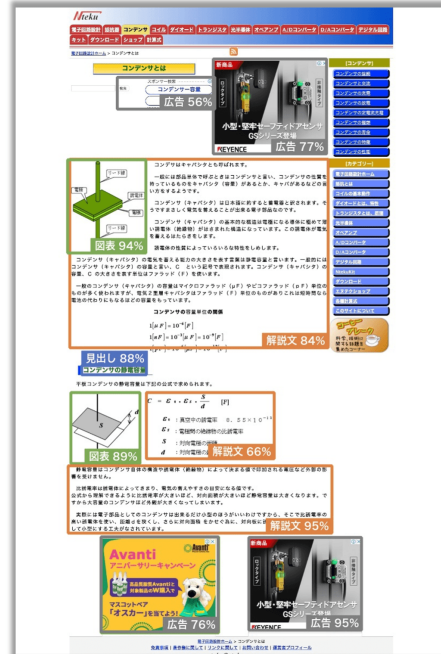


表 6 訓練・開発・評価事例の分野の組み合わせ

番号	訓練事例	開発事例	評価事例
1	線形代数学, 解析学, 力学, 電磁気学, 生物学, 医学, IT	化学	統計学, プログラミング
2	解析学, 統計学, 電磁気学, 化学, 生物学, IT, プログラミング	力学	線形代数学, 医学
3	線形代数学, 統計学, 力学, 電磁気学, 医学, IT, プログラミング	生物学	解析学, 化学
4	線形代数学, 解析学, 電磁気学, 化学, 生物学, 医学, プログラミング	統計学	力学, IT
5	解析学, 統計学, 力学, 化学, 医学, IT, プログラミング	線形代数学	電磁気学, 生物学



(a) 学術分野: 化学



(b) 学術分野: 電磁気学

図 5 ページ構成要素同定モデルの出力例

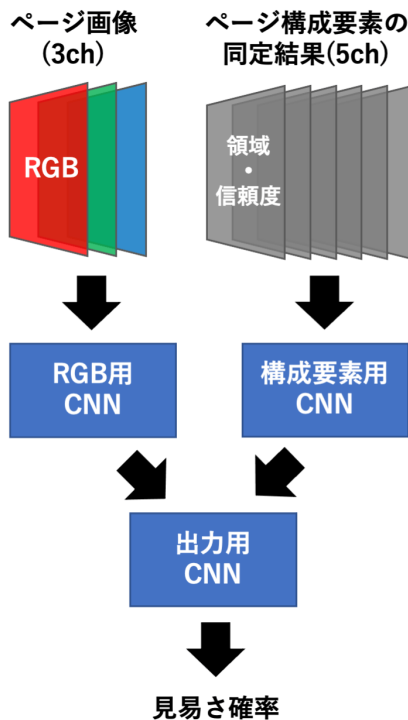


図 6 ページ構成要素を考慮した見易さ自動評価モデル

デル (構成要素考慮モデル) が、ベースラインモデルに比べて高い精度で評価を行えたことが分かる。

## 6 見易さ自動評価結果におけるページ構成要素を考慮した理由提示

### 6.1 見易さ自動評価結果に最も寄与したページ構成要素の特定

本節では、見易さ自動評価結果において、ページ構成要素を考慮して自動評価理由を提示する手法として、見易さ自動評価結果に最も寄与したページ構成要素を特定する手法を提案する。本手法では、4 節において作成した見易さ自動評価モデルの勾配を利用して、「各ページ構成要素ごとの入力チャンネル」の寄与度を算出する (図 8)。まず、入力画像  $X$  に見易さ自動評価モデルを適用し、見易さ評価結果の確率  $y$  を得る。次に、誤差逆伝播法により、見易さ確率  $y$  に対する入力データ  $X$  の勾配  $\frac{\partial y}{\partial X}$  を求める。その後、入力データ  $X$  とその勾配  $\frac{\partial y}{\partial X}$  の中からページ構成要素に関する 5 チャンネル分の要素を取り出し、チャンネルごとに要素積の和を取ることで、各チャンネルの寄与度  $c_i$  ( $i = 1, \dots, 5$ ) を算出する。本論文では、各ページ構成要素のうち、寄与度が正值であり、かつ、最大となる要素を見易さ自動評価結果に最も寄与したページ構成要素であると判定した。

表7 ページ構成要素の同定結果

	見出し	目次	解説文	図表	広告
参照用総数	6,613	356	6,538	1,998	757
Average Precision (AP)	0.83	0.78	0.73	0.78	0.65

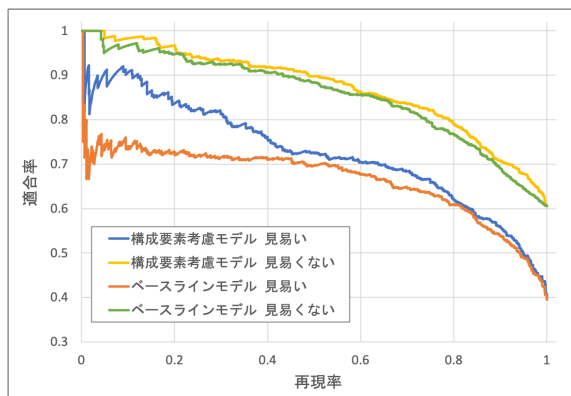


図7 ページ構成要素を考慮した見易さ自動評定モデルの再現率-適合率曲線

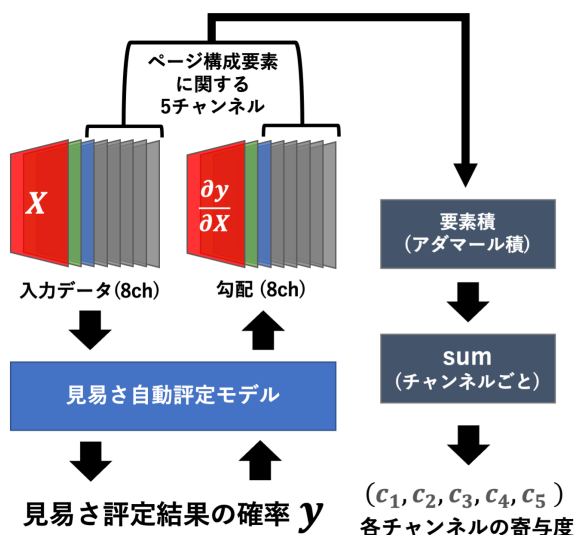


図8 ページ構成要素に関する入力チャンネルの寄与度の算出方法

表8 「見易さ自動評定結果におけるページ構成要素を考慮した理由提示」に対する人手評価基準

評価値	説明
3	「見易さ自動評定結果に最も寄与したページ構成要素」と、「人手による見易さ評定において最も寄与する構成要素」が同一であり、かつ、「見易い/見易くない」の極性も一致している。
2	「見易さ自動評定結果に最も寄与したページ構成要素」は、「人手の見易さ評定において寄与する構成要素」であり、かつ、「見易い/見易くない」の極性も一致しているが、最も寄与する構成要素ではない。
1	「見易さ自動評定結果に最も寄与したページ構成要素」は、「人手の見易さ評定において寄与する構成要素」ではないか、あるいは、「人手の見易さ評定において寄与する構成要素」ではあるが、「見易い/見易くない」の極性が一致しない。

## 6.2 評価手順および評価結果

前節の手順により、学術用語解説ページの全事例 1,945 件のうち 1,425 件において、見易さ自動評定結果に最も寄与したページ構成要素を特定することができた。その中から無作為に選定した 250 件について、表 8 に示す 3 段階の基準に基づき、「見易さ自動評定結果におけるページ構成要素を考慮した理由提示」に対する人手評価を行った。ただしこの評価に先立ち、評価対象の 250 件について、4 節において作成したモデルによってページ構成要素を同定する際の信頼度 (0~1) を参照し、次式で示す「信頼度を考慮した IoU」を算出し、この値が 0.6 未満の事例は評価の対象外とした (2.5 節において人手で付与したページ構成要素の領域の画素集合を  $R$ 、4 節において作成したモデルによってページ構成要素を同定した領域の画素集合を  $P$  とする)。

$$\text{IoU} = \frac{|R \cap P|}{|R \cup P|}$$

$$\text{信頼度を考慮した IoU} = \frac{|R \cap P|}{|R| + |P - (R \cap P)| \times \text{信頼度}}$$

人手評価結果を表 9 に示す。全 250 件のうちの約 60% の理由提示結果が「3」と評価され、特に、見易さ自動評定結果において「見易い」と評定された用語解説ウェブページにおける理由提示結果に対して、「3」の評価割合が高いことが分かった。また、表 9 の結果を、表 7 に示す各ページ構成要素の参照用総数の比率と比較すると、まず、参照用総数が相対的に少ない目次、および、広告については、「見易さ自動評定への影響」も相対的に少ないことが分かる。一方、参照用総数が相対的に多い見出し、解説文、図表については、「見易さ自動評定への影響」が相対的に大きく、そのうち、見出し、および、図表については、「見易い」場合の影響が特に大きいことが分かる。一方、解説文については、「見易い」場合、「見易くない」場合、ともに影響が大きいことが分かる。このうち、特に、見出しについては、参照用総数に対して、影響が大きいと特定された事例数が半減している。これは、「見出しが見易くない」事例の大半において、「見易くない」という評価結果に対して見出しよりも大きい影響を与えた構成要素が別に存在したことが原因である。

ページ構成要素を考慮した理由提示結果とその人手評価例<sup>10 11 12</sup>を図 9 に示す<sup>13</sup>。図 9(a) の事例では、自動評定モデルで「見易い」と判定され、その理由として解説文が提示された。人手による見易さ評定においても解説文が最も「見易い」要素とされたため、「3」と評価した。図 9(b) の事例では、「見易くない」要素として図表が提示された。人手による見易さ評定においても図表が「見易くない」要素とされたが、「見易くない」要素としては、解説文の方がより重大な要因と判定されたため、

10 : 図 9(a): <https://rikigaku-room.com/ennunndou/koushinnennshinnchigai/>

11 : 図 9(b): <http://support.avis.ne.jp/user-support/connect/faq/ip-address>

12 : 図 9(c): [https://wakatech.jp/object\\_oriented](https://wakatech.jp/object_oriented)

13 : 図中において、見易さ自動評定結果に最も寄与したとして特定したページ構成要素の位置をページ画像上の矩形によって示す。

表9 「見易さ自動判定に最も影響したページ構成要素の特定」の人手評価結果 (ページ数)

ページ 構成要素	モデルによる自動判定結果								合計
	見易い				見易くない				
	人手評価値			「信頼度を考慮した IoU」が 0.6 未満 (%)	人手評価値			「信頼度を考慮した IoU」が 0.6 未満 (%)	
	3 (%)	2 (%)	1 (%)		3 (%)	2 (%)	1 (%)		
見出し	39 (15.6)	6 (2.4)	0 (0)	3 (1.2)	1 (0.4)	9 (3.6)	3 (1.2)	2 (0.8)	63 (25.2)
目次	5 (2.0)	1 (0.4)	0 (0)	0 (0)	0 (0)	1 (0.4)	2 (0.8)	0 (0)	9 (3.6)
解説文	42 (16.8)	22 (8.8)	2 (0.8)	3 (1.2)	46 (18.4)	2 (0.8)	0 (0)	5 (2.0)	122 (48.8)
図表	13 (5.2)	0 (0)	1 (0.4)	0 (0)	3 (1.2)	8 (3.2)	10 (4.0)	0 (0)	35 (14.0)
広告	0 (0)	0 (0)	9 (3.6)	1 (0.4)	0 (0)	3 (1.2)	7 (2.8)	1 (0.4)	21 (8.4)
合計	99 (39.6)	29 (11.6)	12 (4.8)	7 (2.8)	50 (20.0)	23 (9.2)	22 (8.8)	8 (3.2)	250 (100)



学術分野：力学  
見易さの人手評価：見易い  
見易さの自動判定確率 (0~1)：0.93  
特定した構成要素：解説文 (見易い)  
(a) 理由提示結果の評価：3



学術分野：IT  
見易さの人手評価：見易くない  
見易さの自動判定確率 (0~1)：0.13  
特定した構成要素：図表 (見易くない)  
(b) 理由提示結果の評価：2



学術分野：プログラミング  
見易さの人手評価：見易い  
見易さの自動判定確率 (0~1)：0.59  
特定した構成要素：広告 (見易い)  
(c) 理由提示結果の評価：1

図9 見易さ自動判定結果におけるページ構成要素を考慮した理由提示に対する評価例 (見易さ自動判定結果に最も寄与するページ構成要素)

「2」と評価した。一方、図9(c)の事例では、人手による見易さ評価では「見易さ」を阻害する要素と判定された広告が、自動判定モデルでは「見易い」理由として提示されたため「1」と評価した。

## 7 おわりに

本論文では、学術用語を解説するウェブページ群を対象として、用語解説ページの画像データに加えて、ページ構成要素 (見出し、目次、解説文、図表、広告) の領域情報を考慮することにより、見易さ自動判定モデルの性能を改善するとともに、評定理由の具体化が行えることを示した。今後の課題としては、ページ構成要素を考慮した評定理由の提示内容について、より詳細な分析・評価を行うことが挙げられる。また、本論文の見易さ自動判定システムと併用して、文献[8]における「文章の分かり易さ」自動判定システム、および、文献[3]における「学

術用語解説ページの充実度・対象読者の習熟度」自動判定システムを導入することにより、学術用語解説ページの「学び易さ」を分析するシステムを作成することが挙げられる。

## 文 献

- [1] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proc. 29th CVPR*, pp. 770–778, 2016.
- [2] T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft COCO: Common objects in context. In *Proc. 13th ECCV*, pp. 740–755, 2014.
- [3] 大賀悠平, 岡田心太郎, 曾田耕生, 宇津呂武仁, 河田容英. 学術用語解説ウェブページの充実度・対象読者の習熟度評定データセットの作成. 第 13 回 DEIM フォーラム論文集, 2021.
- [4] S. Okada, C. Hirohara, K. Kawaguchi, K. Soda, T. Utsuro, Y. Kawada, and N. Kando. Identifying factors of visual intelligibility of Web pages explaining academic concepts. In *Proc. DL4Ed*, 2019.
- [5] 岡田心太郎, 塩川隼人, 韓炳材, 廣花智通, 宇津呂武仁, 河田容英, 神門典子. 深層学習による学術用語解説ウェブページの見易さ自動判定結果の理由提示. 第 11 回 DEIM フォーラム論文集, 2019.
- [6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In *Proc. 29th CVPR*, pp. 779–788, 2016.
- [7] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proc. 16th ICCV*, pp. 618–626, 2017.
- [8] 塩川隼人, 岡田心太郎, 韓炳材, 廣花智通, 宇津呂武仁, 河田容英, 神門典子. 深層学習を用いた学術用語解説ウェブページの分かり易さ・見易さの自動判定. 第 11 回 DEIM フォーラム論文集, 2019.