

# 遠隔歯科診療に向けた 自撮り口腔内写真の利用可否判定手法

片山 洋平<sup>1</sup> 松井 智一<sup>1</sup> 松田 裕貴<sup>1</sup> 諏訪 博彦<sup>1,2</sup> 安本 慶一<sup>1</sup> 上田 啓太<sup>3</sup> 竹山 旭<sup>3</sup>

**概要：**口腔ケアは健康寿命の延伸に重要である。一方で、自身で口腔内の衛生状態を判断するのは難しく、歯科医師の支援が必要である。近年のオンライン診療の普及に伴い、歯科においても口腔内診断に関する研究がなされている。歯科のオンライン診療では、患者がスマートフォンで自身の口腔内を撮影し、歯科医師が遠隔で診療を行うことが望まれる。一方で患者にとって診断利用可能な画像かを判断することは困難であり、撮り直しが多発した結果としてオンライン診療を断念するおそれがある。そこで、本研究では患者の5枚撮影法によるスマートフォンでの自撮り撮影を想定し、口腔領域の検出および診断利用可否判定モデルを提案する。

**キーワード：**口腔健康管理、口腔内写真、オンライン診療支援

**Abstract:** Oral care is important for extending healthy life expectancy. On the other hand, it is difficult to manage oral hygiene on one's own and requires the support of a dentist. In recent years, with the spread of remote treatment, research has been conducted on online oral diagnosis. In online oral diagnosis, patients take pictures of their own oral photo with a smartphone, and the dentist performs the examination remotely. It is difficult for the patient to determine whether the images are diagnostic and usable, and there is a risk that the images may have to be retaken frequently. In this study, we propose a method for detecting oral regions and determining whether or not they can be used for diagnosis, based on the assumption that patients take selfies with their smartphones using the 5 angles image method.

**Keywords:** Oral health care, Oral cavity photo, Online diagnosis support system

## 1. はじめに

口腔ケアは健康寿命の延伸に重要なことが近年わかってきてている。また、厚生労働省の提唱する「健康21」では歯科疾患は自覚症状が伴わず発生することが多く、早期診断・定期的な検診をすることが重要と記されている[1]。全身の健康を保つためには歯の状態を清潔に維持することは重要である一方で、自分で衛生状態を判断するのは難しく、歯科医師や歯科衛生士の支援が必要である。さらに、新型コロナウイルス感染症などの影響を受け、実際に歯科医院に赴き歯科医師の診断を受けることへのハードルが高まっている。このような背景から、歯科においてもオンライン診療の需要が高まっている。日本国内でも政府が時限的に初診からのオンライン診療を解禁している[2]。また、

第44回国家戦略特別諮問会議ではコロナ収束後もオンライン診療をあらゆる分野で積極的に活用する検討をしているとの報告がされた[3]。一般的に歯科の診療は、歯科医師が目視で行われるため、歯科のオンライン診療では、口腔内の画像で口腔内の状況を伝えることが最も効果的といえる。そのためには、写真撮影や動画撮影などを患者自らが行う必要があるものの、奥歯や咬合面の撮影は患者が直接目視することができないため、容易ではない。また、患者は歯科分野に関する専門知識は備えていないため、撮影された口腔内写真が歯科医師が診療を行う上で十分な画像であるか否か判断できない。場合によっては、撮り直しが発生し、オンライン診療を離脱してしまうという機会損失につながることが想定される。したがって、歯科のオンライン診療を実現するためには、歯科医師が診断を行うことにあたって十分な画像を患者自身が撮影することを支援するシステムが必要であると考えられる。

遠隔で歯科診療を行う際には、複数の角度から撮影された口腔内写真が必要となる。熊谷ら[4]は歯科医師に対し

<sup>1</sup> 奈良先端科学技術大学院大学

Nara Institute of Science and Technology

<sup>2</sup> 理化学研究所 革新知能統合研究センター

RIKEN Center for Advanced Intelligence Project

<sup>3</sup> 株式会社 NOVENINE

NOVENINE inc.



図 1 5枚撮影法 [4] により得られる画像群

て診断や治療経過の観察のための口腔内規格写真を提案している。これらの規格写真では5枚撮影法や12枚撮影法といったように写真組みを定義している。各撮影法では、側面や咬合面といった多様な角度からの撮影を組み合わせることで、口腔内の状態を網羅的に写真として保存する。本研究では、患者の負担を最小限にするため、図1に示すような5枚撮影法を対象として自撮りで撮影する際の支援を行う。歯科の臨床現場においては、口を大きく開けるため、口角鉤を用いて撮影を行っており、道具を使わずに患者自身で十分に口を広げるのは困難である。アングルワイダーと呼ばれる開口器具は安価に入手可能であり、このアングルワイダーを用いることで、不慣れな患者自身でも十分に広がった口腔内の撮影ができるようになると考えた。

本研究では、患者が5枚撮影法の各角度に関して撮影時にスマートフォンを口にかざすことで、診断利用可能な画像を自動撮影する手法を提案する。提案手法ではアングルワイダーにマーカを取り付け自動的に口腔領域の位置検出を行い、深層学習により歯科診断可否判定を行うモデルを構築する。提案手法は患者に対して撮影の補助、撮り直しの減少を提供し、歯科医師に対しては真に診断利用可能な画像のみを提供することが可能となる。したがって、本手法により患者・歯科医師のいずれに対しても負荷の少ないシステムを提供することができる。

構築した診断利用可否判定モデルの性能評価実験では、21人の口腔内画像を収集し、構築した診断利用可否判定モデルの精度を検証した。

## 2. 関連研究

### 2.1 歯科診療に関する研究

近年、スマートフォンの普及や高画質化の影響により、患者の自撮り写真を口腔内疾患の診断やスクリーニング検査に用いる研究が為されている。LiangらのOralCam [5]ではスマートフォンのインカメラで撮影された口腔内画像に対して、深層学習により疾患の有無や場所を自動で画像認識するシステムを提案している。東北大学とNTTドコモは共同でスマートフォンで自撮りされた正面観の口腔画像から歯周病の自動検知の研究を行なっている [6]。また、ライオン株式会社は自撮り画像の舌に映る舌苔から口臭の

リスク判定を行なっている [7]。このように実際に患者が撮影した画像をもとに、ある特定の口腔疾患やその予兆を自動判定する研究はいくつか存在する [8,9]。しかし、AIによる自動画像診断は研究としては存在するものの、あくまでも歯科医師等の診断をサポートするものであり、実用化には至っていない。またこれらの研究では「スマートフォンを用いた口腔内の撮影」が正しく行われることを前提としているが、自身の口腔内を自ら撮影することは日常的な撮影とは大きく異なり、正しく撮影することは非常に困難といえる。歯科の臨床現場にて診断等に活用されている5枚撮影法などに基づく撮影は、奥歯の側面の撮影なども含まれ、患者が撮影することは容易ではない。また正しく撮影されていない画像に基づく診断は判断のあやまりを招くため、歯科医師は適切な画像を共有してもらうために、何度も患者と画像をやり取りする恐れがある。現時点では、歯科医師に診断を行ってもらう上で十分な口腔内画像を患者が容易に撮影できるように支援するシステムではなく、患者・歯科医師・ネットワークの全てにおいて負荷の軽い仕組みが望まれている。

### 2.2 自撮り支援に関する研究

自撮り撮影を補助しているシステムとして、oh my teeth [10]では歯列矯正を遠隔で受ける患者のため、スマートフォンでの撮影時の画面に歯形を表示し、手振れ補正を行っている。これにより歯科医師が経過を確認できる写真の撮影の補助を行っている。しかし、患者が画面を確認しながら正しい角度に近づけて行く必要があるため咬合面のように撮影時にスマートフォンの画面が見えない場合には適していない。より一般的な自撮り撮影の補助機能としてスマイルシャッター [11]がある。これは画像認識により被撮影者の笑顔を検出することで、最も適したタイミングで自動的にシャッターを切る機能としてデジタルカメラやスマートフォンに搭載されている。本研究でも同様に、患者がかざすスマートフォンから歯科医師の診断利用可能な画像をリアルタイムに検出し撮影を行う手法を提案する。

## 3. 診断利用可否判定手法

本章では5枚撮影法を患者が撮影する支援をするため、

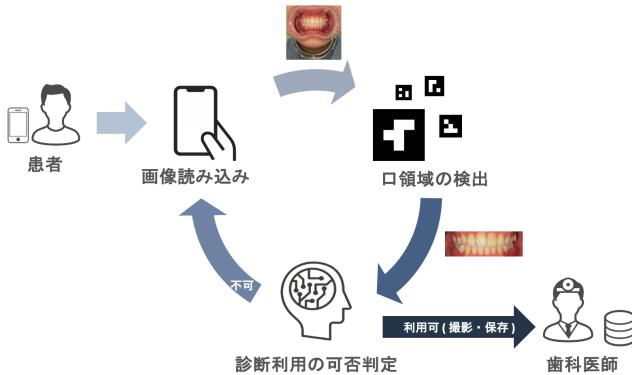


図 2 口腔内写真撮影フロー

スマートフォン上で各角度に関して診断利用可能な画像か否かをリアルタイムに判断する手法を提案する。

### 3.1 口腔内写真撮影フロー

図 2 に提案する口腔内写真撮影フローを示す。初めに患者は 5 枚撮影法の各角度で、背面カメラを自分の口にかざす。初めにマーカーを用いて口領域の検出を行う。続いて、検出された口領域に対して深層学習モデルを用いて歯科医師が診断利用可能な画像かを判定する。ここで診断利用可能とされた画像が現れた際にシャッターを切り、データベース上に保存する。これにより、患者は短時間で口腔内の画像を撮影し、歯科医師は必要最小限の写真から診断を行うことが可能となる。写真的解像度は口腔内診断に影響すると考えられるため、スマートフォンの背面カメラで取得する。

### 3.2 口領域の検出

リアルタイムに口腔内写真が診断利用可否か判定するため予め口領域を検出する。口領域の Bounding Box を抽出することは、後の診断利用の可否判定モデルの性能を向上させるためにも重要である。昨今では高速かつ高精度なオブジェクト検出モデルとして YOLOv3 [12] や YOLOv3-tiny [13] が提案されている。これらのモデルはファインチューニングすることにより様々なタスクに適用できる一方で、モデル構成が複雑でありスマートフォン上で十分な速度や精度を保って動作させることは難しい。そこで、本研究ではアングルワイダーにマーカーを設置し、マーカーの位置から口領域の検出を行う。マーカーには AR マーカーを用いる。早川らの研究 [14] では、複数の AR マーカーを歯ブラシの後端に付与してスマートフォンのカメラで位置推定を行い歯磨き動作を検出している。このように、AR マーカーは従来の拡張現実を提示する目的に限らず位置を取得する用途でも用いられることがある。図 3 に本研究でアングルワイダーに付与した AR マーカー

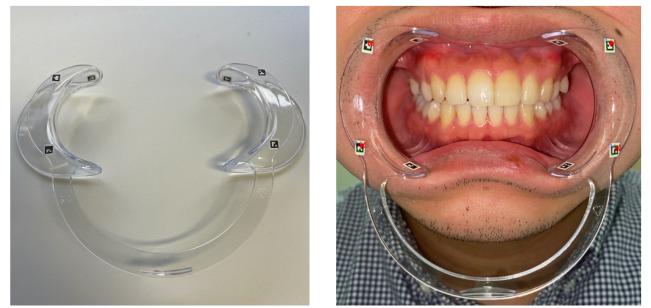


図 3 AR マーカーを付与したアングルワイダーと正面観の検出結果

と、検出結果を示す。マーカーは 5 枚撮影法の各角度において少なくとも 2 枚が映り込むように 8 枚配置する。検出は OpenCV の ArUco ライブラリ [15] を用いてスマートフォン上で各 AR マーカーを識別するための ID および位置情報を取得する。正面観はアングルワイダーの表面に付いている 4 枚から口領域を検出し、左右側方面観および上下顎咬合面観では口領域の縦横比率を定め、検出される 2 枚のマーカーの距離から抽出する口領域を算出する。ここで検出された口領域画像を、続く診断利用可否判定モデルの入力とする。

### 3.3 診断利用の可否判定

患者が撮影した口腔内写真が診断に使えるか否かは、歯・歯肉の見え方や照明および画像の解像度など様々な要因が影響する。また、歯科医師は診療のため全ての歯について確認する必要がある。そこで本研究では、深層学習モデルである CNN [16] を用いて歯科医師が診断利用可能な口腔画像か否かを判定するモデルを構築する。スマートフォン上の動作を行うため、2 層の畠み込み・プーリング層および 1 層の全結合層から成る 3 層で構成する。また、モデル学習時の過学習の抑制、学習の最適化を行うため、中間層にバッチ正規化層、ドロップアウト層を追加する。なお、CNN のユニット数などのパラメータは経験的に設定している。利用可否の正解ラベルには、歯科医師が 5 枚撮影法の各角度について撮影要件を定め、歯科医師と歯科助手により判定を行う。

CNN の学習において、歯列は個人差があるためデータセットを擬似的に拡張することが有効であると考えられる。深層学習においてデータ拡張は有効性であるとされている [17] が、本研究では拡張手法によっては診断利用可能であった画像が利用不可になる可能性があり、その都度正解ラベルを設定するのはコストが高い。したがって各角度におけるデータ拡張手法を提案する。

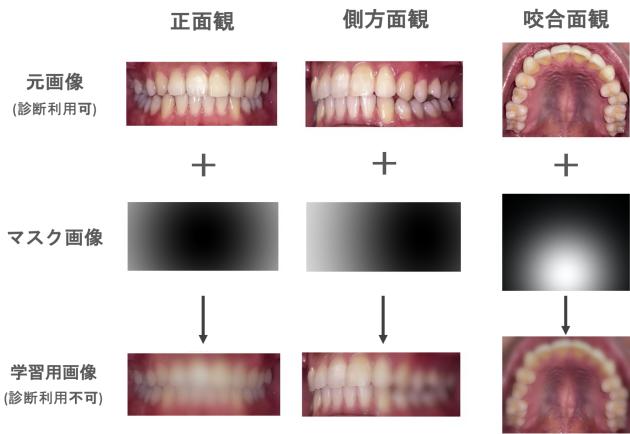


図 4 ガウシアンフィルターを用いた学習用のぼかし画像の生成

### 正面観・咬合面観のデータ拡張

正面観および上下顎咬合面観で利用可能な画像を拡張する方法として左右反転を行う。また、利用可能な画像から、不可な画像を拡張する方法としてぼかし・クロップ処理を行う。ぼかし処理の例を図 4 に示す。ここでマスク画像の黒色部分は各角度において歯科医師が注目する場所を示す。この注目する部分にガウシアンフィルタをもちいてぼかし処理を加えることで、利用不可な画像を擬似的に生成する。またクロップ処理も同様に注目する場所が全て入らないよう左右にクロップを行い利用不可な画像を生成する。

### 側方面観のデータ拡張

左右側方面観では歯科医師は奥歯を注目しているため、そのまま反転画像を用いることはできない。そこで反対側の側方面観の利用可能画像を反転したものラベルを変更せずに用いる。ぼかし処理では奥歯にぼかしを入れるマスク画像を作成し、利用不可な画像を擬似定に生成する。また、クロップでは奥歯を含むように切り出した際には利用可能に、奥歯が映り込まないように切り出した際には利用不可になるようにラベルを修正する。

## 4. 診断可否判定モデルの評価実験

提案する診断利用可否判定モデルの性能を評価するための実験を実施する。

### 4.1 実験内容

評価実験では、5枚撮影法により男女21名の口腔内画像を収集する。8名は自宅で自撮りにより撮影を行い、13名は歯科医院にて他撮りにより撮影された。すべての画像はスマートフォンの背面カメラを用いて撮影を行う。また、診断利用可否のラベルは歯科医師1名と歯科衛生士1名がそれぞれ行った。2名の判定が異なる際には再度確認し、双方の同意が得られる判定を取得した。実験結果にはこの2名の判定の一一致率についても表記する。

診断利用可否判定モデルに関する評価を行うため、本実

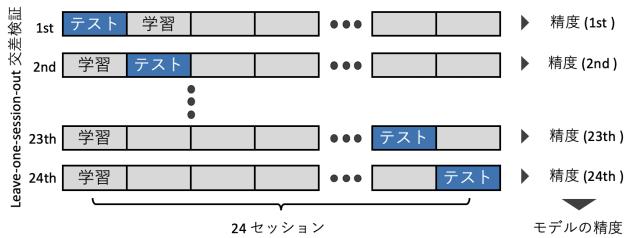


図 5 各角度での交差検証の模式図

験では手動により口領域の切り出しを行い、それらの画像を用いた。ここで明らかに口領域が判断できない画像に関してはデータセットから除外する。同じ環境で撮影されたデータ同士を学習・テストデータに含めることは実際には不可能なため、図 5 に示すような、Leave-one-session-out 交差検証を行い適合率、再現率、F 値により評価する。1 セッションでは、すべての被験者が少なくとも 1 回は 5 枚撮影法における各角度で、口腔内を撮影する。本実験では、学習・評価に利用できる十分なデータ数を確保するため、21 名の被験者のうち 3 名が 2 度セッションを行っていることから、合計で 24 セッションの交差検証によりモデルを評価する。

### 4.2 実験結果

表 1 に本実験で用いたデータ、ならびに提案手法による推定結果を示す。また、図 6 に各角度における推定結果の混同行列を表示する。結果より、正面観が最も高い 84.3% の F 値で判定可能であり、左側側方面観が最も低い 67.6% の F 値で判定可能であった。したがって、撮影方向により、判定精度が異なることがわかった。

全角度における精度として、適合率が高く、再現率が低い結果となった。実際のオンライン診療では、患者の所持するスマートフォンで連続的に画像を読み込み、診断可能が高いスコアで推定された画像のみをサーバへアップロードすることが想定される。一般に、適合率と再現率はトレードオフの関係である。適合率を向上させたシステムは、患者に長時間の撮影を求め、再現率を向上させたシステムは、歯科医師に不要な画像を確認させることとなる。ここで、患者はスマートフォンで連続的に画像を取得するため、診断利用可能な画像が取得できるまで撮影することはある程度の許容が可能である。一方で、医師は受信した画像について 1 枚ずつ確認を行うため、そこで判定不可として却下するリスクを考慮すると、再現率よりも確実に診断に利用可能な確率を示す適合率の向上が求められる。また、適合率を向上させることは、不必要的画像データを判定用データプールへアップロードすることを防ぎ、トラフィックを軽減することにも繋がる。

表 1 データセットの内訳と交差検証結果

5 枚撮影法	撮影枚数		被撮影者数	平均枚数	最少枚数	最多枚数	正解ラベル	交差検証		
	利用可	利用不可						一致率	適合率	再現率
正面観	105 枚	94 枚	21 人	8.3 枚	4 枚	29 枚	95.3%	86.9%	81.9%	84.3%
左側側面観	120 枚	90 枚	21 人	8.7 枚	3 枚	25 枚	88.8%	80.5%	58.3%	67.6%
右側側面観	126 枚	83 枚	21 人	8.7 枚	5 枚	25 枚	88.5%	87.4%	65.9%	75.1%
上顎咬合面観	98 枚	80 枚	21 人	7.7 枚	1 枚	27 枚	85.6%	82.5%	67.3%	74.2%
下顎咬合面観	105 枚	95 枚	21 人	8.3 枚	2 枚	25 枚	89.7%	82.5%	76.2%	79.2%

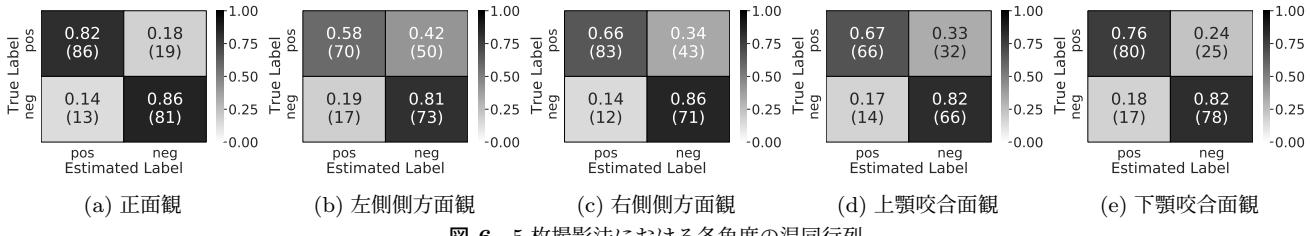


図 6 5 枚撮影法における各角度の混同行列

### 4.3 考察

診断利用可否判定モデルの評価では正面観が最も高い精度を得た。正面観は利用可否のアノテーションに関して最も一致率が高く、どのような画像が利用可能か定義が明確であることが確認できる。一方で他の角度に関してはアノテーションの一一致率も 90% 程度であり、診断利用の可否判断における個人差が確認できる。このことは側面や咬合面の精度が正面観と比較して、低下している原因の 1 つとして考えられる。本研究において、出力される診断利用可能な画像は多くの歯科医師にとって利用可能であることが望ましい。アノテーションを行う歯科医師の人数を増やし、多数決によるアンサンブル学習 [18] を行うことでこれらの課題を解決可能であると考えられる。

今回のデータセットでは、前歯が抜けている患者が 1 名含まれた。この患者から得られた画像は、他の患者の歯型と大きく離れているため、学習におけるノイズになったと考えられる。他にも治療に FCK(銀歯)を利用している場合や歯列が他の患者と比べて大きく異なる患者のように個人差が大きい画像が存在し、精度に影響した。これらの精度低下に関しては被験者数を増やすことにより改善可能であると考えられる。

今後は、AR マーカーによる画像切り出しから構築した診断利用可否判定モデルまでのパイプラインを通じた評価により、実際に患者が自宅で自撮りを撮影する際の運用における利用可否を評価する必要がある。

### 5. まとめ

本論文では、歯科のオンライン診療に向けた自撮り口腔内写真の利用可否判定手法の提案を行った。提案手法ではマーカーによる口領域の検出、CNN による診断利用可否判定のモデルの構築を行った。利用可否判定モデルの性能を評価するため、21 名を対象に 5 枚撮影法にもとづいたデータセットを収集し、Leave-one-session-out 交差検証に

よりモデルの評価を行った。その結果、F 値では正面観が最も高く 84.3% となり、左側側面観が最も低く 67.6% となった。この要因として、正面観では正解ラベルの一一致率が最も高いことから利用可否の判断を行う上で個人差が少ないためであると考えられる。今後の課題としては、被験者数および各被験者のデータ数を拡張し学習することで汎用的な可否判断モデルを構築するとともに、複数の歯科医師のアノテーションを取得することにより多数の歯科医師が診療を行う上で診断利用可能であるというラベルを更新する必要がある。また、スマートフォン上でマーカー検出から口領域の切り出し、利用可否判断を行った際の認識精度および認識速度を評価する必要がある。

謝辞 データ収集にあたり協力いただきました株式会社 NOVENINE のみなさま、画像のアノテーションにあたりご協力いただきました歯科医師の竹山先生、歯科衛生士の川口様に心より感謝申し上げます。

### 参考文献

- [1] 厚生労働省. 健康日本 21(歯の健康). 2000. [https://www.mhlw.go.jp/www1/topics/kenko21\\_11/pdf/b6.pdf](https://www.mhlw.go.jp/www1/topics/kenko21_11/pdf/b6.pdf) (accessed on 18 Feb. 2021).
- [2] 厚生労働省医政局医事課. 事務連絡「新型コロナウイルス感染症の拡大に際しての電話や情報通信機器を用いた診療等の時限的・特例的な取扱いについて」. 2020.
- [3] 国家戦略特区. 第 44 回 国家戦略特別諮問会議. 2020.
- [4] 熊谷崇, 熊谷ふじ子, 鈴木昇一. 新口腔内写真的撮り方. 東京 医歯学出版, 2007.
- [5] Yuan Liang, Hsuan Wei Fan, Zhujun Fang, Leiyi Miao, Wen Li, Xuan Zhang, Weibin Sun, Kun Wang, Lei He, and Xiang'Anthony' Chen. Oralcam: Enabling self-examination and awareness of oral health using a smartphone camera. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020.
- [6] 東北大学大学院歯学研究科プレスリリース. 東北大学とドコモ、歯周病発見 A I の共同研究を開始, 2019. <https://www.tohoku.ac.jp/japanese/2019/>

- 02/press-20190221-02-docomo.html (accessed on 18 Feb. 2021).
- [7] ライオン株式会社. Repero, 2018. <https://repero.lion.co.jp/> (accessed on 18 Feb. 2021).
  - [8] Guy Tobias and Assaf B Spanier. Modified gingival index (mgi) classification using dental selfies. *Applied Sciences*, Vol. 10, No. 24, p. 8923, 2020.
  - [9] 歯っびー. 歯っびー, 2019. <https://plaque-detection.jimdofree.com/> (accessed on 18 Feb. 2021).
  - [10] oh my teeth. oh my teeth, 2020. <https://prtmes.jp/main/html/rd/p/000000006.000048884.html> (accessed on 18 Feb. 2021).
  - [11] ソニー株式会社. スマイルシャッター. 2007. [https://www.sony.jp/cyber-shot/products/DSC-T70/feature\\_1.html](https://www.sony.jp/cyber-shot/products/DSC-T70/feature_1.html) (accessed on 18 Feb. 2021).
  - [12] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.
  - [13] Pranav Adarsh, Pratibha Rathi, and Manoj Kumar. Yolo v3-tiny: Object detection and recognition using one stage improved model. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICAACCS)*, pp. 687–694. IEEE, 2020.
  - [14] 早川修平, 吉田孝博, 半谷精一郎ほか. Ar マーカを装着した歯ブラシによる歯磨き位置測定手法の検討. 研究報告コンピュータビジョンヒューマニメティア (CVIM), Vol. 2019, No. 5, pp. 1–4, 2019.
  - [15] Sergio Garrido-Jurado, Rafael Muñoz-Salinas, Francisco José Madrid-Cuevas, and Manuel Jesús Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, Vol. 47, No. 6, pp. 2280–2292, 2014.
  - [16] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278–2324, 1998.
  - [17] Luke Taylor and Geoff Nitschke. Improving deep learning using generic data augmentation. *arXiv preprint arXiv:1708.06020*, 2017.
  - [18] Maryam Sabzevari, Gonzalo Martínez-Muñoz, and Alberto Suárez. Vote-boosting ensembles. *Pattern Recognition*, Vol. 83, pp. 119–133, 2018.