



Data Article

Varying views of maxillary and mandibular aspects of teeth: A dataset



Shweta Chaudhary^{a,*}, Preetam Shah^a, Priyanka Paygude^b,
Shwetambari Chiwhane^c, Pratibha Mahajan^d, Prashant Chavan^b,
Manisha Kasar^b

^a Bharati Vidyapeeth Deemed to be University Dental College and Hospital, Katraj, Pune, India

^b Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India

^c Symbiosis Institute of Technology, Symbiosis International University, Pune, India

^d Department of Artificial Intelligence, Vishwakarma University, Pune, India

ARTICLE INFO

Article history:

Received 3 June 2024

Revised 30 June 2024

Accepted 18 July 2024

Available online 25 July 2024

Dataset link: [Teeth or Dental Image Dataset \(Original data\)](#)

Keywords:

Dental photography

Maxillary view

Mandibular view

Mixed dentition

Non-cariou teeth

Occlusal view

Permanent dentition

Primary dentition

ABSTRACT

Real teeth or dental image datasets are a valuable resource that is transforming the field of dentistry by enabling automation, improving diagnostics and accelerating research and development. This article presents a comprehensive dataset containing 9,562 images of healthy teeth (non-cariou) from children aged 1 to 14 years. The images capture different views of the teeth, including maxillary (upper) and mandibular (lower) arches, front, right, left, and occlusal (biting surface) views. These images are stored under eight subcategories in the Mendeley repository, a platform for research data. The potential application of this dataset involves using machine learning to analyze the dental condition. This could provide a faster analysis and facilitate remote assessment of dental conditions in underserved areas. Overall, this dataset seems like a promising tool for advancing dental care through the power of machine learning.

© 2024 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>)

* Corresponding author.

E-mail address: shweta.chaudhary@bharativedyapeeth.edu (S. Chaudhary).

Specifications Table

Specific details of the dataset

Subject	Applied Machine Learning, Artificial Intelligence, Big Data Analytics, Dentistry
Specific subject area	Varying aspects of teeth in the maxillary and mandibular arches.
Type of data	Images
Data collection	There are 9562 pictures of noncarious teeth from the mandibular and maxillary arches of children aged 1 to 14 years in various angles. The views primarily taken are Maxillary front, right, left and occlusal views and mandibular front, right, left and occlusal views, a total of eight categories based on the view captured. Images taken under controlled conditions and uploaded to the Mendeley repository make up each category of the dataset. This dataset has great potential to contribute to advancements in dentistry using machine learning techniques.
Data source location	Bharati Vidyapeeth English Medium School and Bharati Vidyapeeth Deemed to be University Dental College and Hospital, Katraj, Dhanakawadi, Pune- 411043. Latitude 18.456610 or 180 27' 24'' north Longitude 73.85630 or 73051' 23'' east
Data accessibility	Repository name: data.mendeley.com Data identification number: 10.17632/6zsnhrds9t.1 Direct URL to data: https://data.mendeley.com/datasets/6zsnhrds9t/1

1. Value of the Data

- This dataset is of value as it provides comprehensive information of all the sets of teeth; i.e. primary, mixed and permanent dentitions in the form of images. A very wide age range from 1 to 14 years is selected which includes all the children. 4 aspects of the dentition are taken in each arch which enables a broad scope of assessment. Children with systemic diseases, growth and developmental disturbances and history of orthodontic treatment were excluded from the dataset.
- A dental assessment dataset provides a practical example of how image processing can be applied in a real-world medical setting. This dataset would be a rich resource for teaching image processing techniques like segmentation, detection and feature extraction.
- This dataset can be used to train and assist AI models in the following:
 - i. Identification of types of teeth (Incisors, canines, premolars, molars) and primary, mixed or permanent dentition alongwith consideration of missing teeth.
 - ii. Identification of any developmental disturbances like amelogenesis imperfecta, dentinogenesis imperfecta, etc.
 - iii. Differentiation between diseased and non-diseased tooth like detection of caries, periodontal problems, etc.
- This dataset serves as a valuable benchmark for evaluating and comparing the performance of various computer vision algorithms in teeth recognition tasks. This allows researchers to evaluate and improve their own models.
- The dataset can be used to develop dental apps with the ability for identification of teeth, cavity detection and dental malalignment.
- This can lead to innovative approaches to dental assessment that might not have been explored within the traditional confines of dentistry alone.

2. Background

Early Childhood Caries (ECC) is the most common pediatric disease globally and is associated with health disparities for children from disadvantaged backgrounds. If identified early on, ECC is avoidable and reversible. However, a lot of kids have obstacles getting dental treatment because

of ignorance or occasionally negligence. A caries detection system that can be used at home may make dental care more accessible to people from all socioeconomic backgrounds and contribute to a significant increase in the number of parents and guardians seeking treatment for their children.

It is the need of the hour to develop artificial intelligence (AI) in such a way that AI powered technology would be used as a diagnostic test (index test) to detect dental caries and provide a caries risk assessment using children's photographs.

To facilitate development of artificial intelligence based diagnostic tests, collection of data set of noncarious teeth as a first phase and later carious teeth as a second phase was planned.

3. Data Description

This [1] is a dataset of 9562 pictures of children's noncarious teeth from the maxillary and mandibular arches in various angles, ranging in age from 1 to 14. Table 1 furnishes specifications of the dataset in a detailed manner. The views primarily taken are Maxillary/ Upper Front (1075 images), Maxillary/ Upper Right (1143 images), Maxillary/ Upper Left (1289 images) and Maxillary/ Upper Occlusal (1140 images) and Mandibular/ Lower Front (926 images), Mandibular/ Lower Right (1401 images), Mandibular/ Lower Left (1566 images) and Mandibular/ Lower

Table 1
Dataset folders with their specifications.

Folder Sr. No	Folder Title	View	Description of View	Number of Images	Resolution In pixels
1.	Upper Front View	Maxillary Arch View	Includes all the Labial and Buccal surfaces of Maxillary Central Incisors to Canines of both right and left sides	1075	1280×550
2.	Upper Right View	Maxillary Arch View	Includes all the Labial and Buccal surfaces of Distal half of Maxillary Right Canine till the last erupted tooth on the right side of the Maxillary arch	1143	1280×550
3.	Upper Left View	Maxillary Arch View	Includes all the Labial and Buccal surfaces of Distal half of Maxillary Left Canine till the last erupted tooth on the left side of the Maxillary arch	1289	1280×550
4.	Upper Occlusal View	Maxillary Arch View	Includes Occlusal and Incisal surfaces of all the teeth in the Maxillary arch	1140	1024×768
5.	Lower Front View	Mandibular Arch View	Includes all the Labial and Buccal surfaces of Mandibular Central Incisors to Canines of both right and left sides	926	1280×550
6.	Lower Right View	Mandibular Arch View	Includes all the Labial and Buccal surfaces of Distal half of Mandibular Right Canine till the last erupted tooth on the right side of the Mandibular arch	1401	1280×550
7.	Lower Left View	Mandibular Arch View	Include all the Labial and Buccal surfaces of Distal half of Mandibular Left Canine till the last erupted tooth on the left side of the Mandibular arch	1566	1280×550
8.	Lower Occlusal View	Mandibular Arch View	Includes Occlusal and Incisal surfaces of all the teeth in the Mandibular arch	1022	1024×768

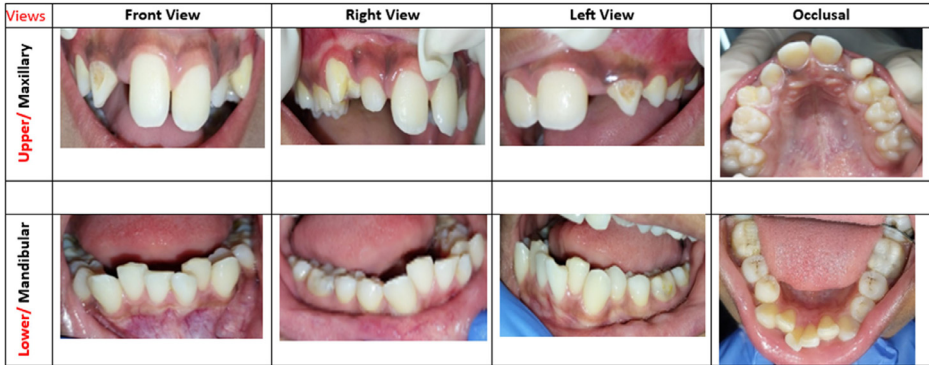


Fig. 1. Sample images of the eight views of the teeth.

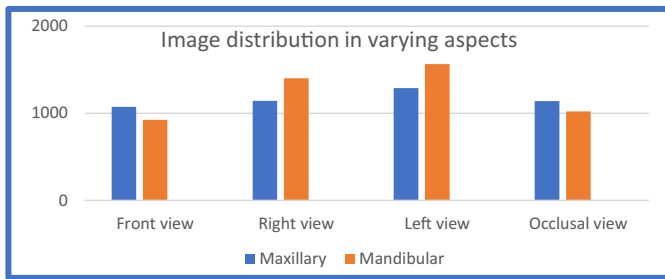


Fig. 2. Image distribution of the Teeth Dataset.

Occlusal (1022 images) as represented in Figs. 1 and 2. Images taken under controlled conditions and uploaded to the Mendeley repository are included in each category of the dataset. This dataset is specifically designed for research in dental assessment using machine learning and computer vision techniques.

The images focus solely on teeth and are of standard resolution of 1280×550 pixels for front, right and left views and 1024×768 pixels for occlusal views. The dataset includes images taken under different lighting conditions, ensuring the model can identify teeth regardless of light levels. The teeth are photographed from eight different angles. This variety helps the model recognize teeth even when they’re not perfectly straight. This diversity in the dataset is designed to make the model more robust in real-world situations where teeth might be seen from various angles, under different lighting and with different backgrounds.

4. Experimental Design, Materials and Methods

4.1. Experimental design

All the images were photographed with a Smartphone (iPhone 152,556×1179 pixels resolution at 460ppi with a 48MP main camera: 26 mm, f/1.6 aperture) after tooth cleaning and drying. Using intraoral mirrors that had been warmed before being positioned in the oral cavity to avoid condensation on the mirror surface, the occlusal surfaces of every tooth were indirectly imaged. These mirrors were used for 2 views, viz: Maxillary/ Upper Occlusal and Mandibular/ Lower Occlusal. The intraoral photographic mirror was placed with its posterior end resting behind the last erupted tooth in the arch and its anterior end extending beyond the Incisor teeth,

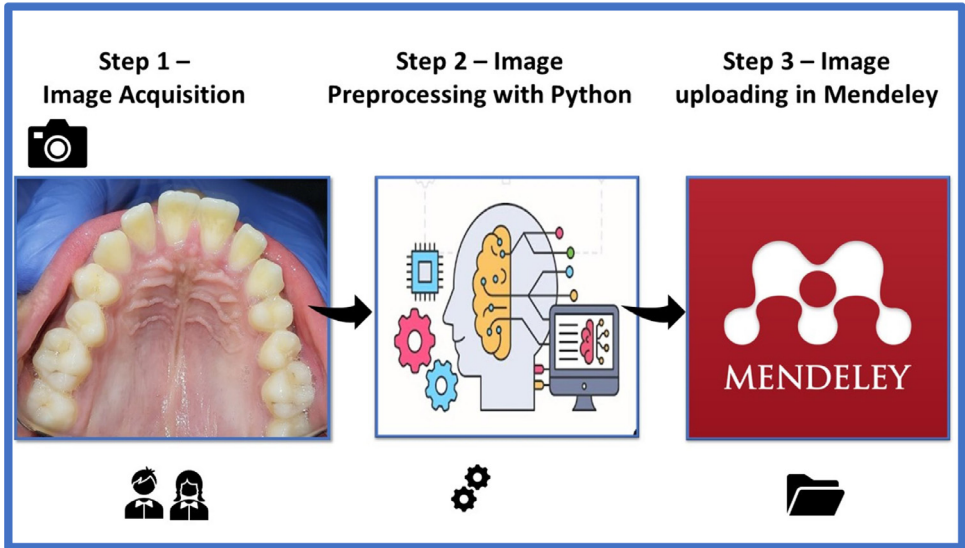


Fig. 3. Data acquisition process.

thus its reflecting surface covering occlusal/ incisal surfaces of all teeth in the arch. For all photographs, distance between the smartphone and the teeth was 5–7 cm with the lens in the same horizontal plane as the teeth. Eight prominent aspects of the teeth were captured. Fig. 2 shows image distribution of the dataset.


The data acquisition process consisted of the following steps, as summarized below in Fig. 3:

- i. Step 1: Collection of images by clicking under controlled conditions as mentioned in Table 1.
- ii. Step 2: Editing of images using Microsoft Windows Snipping Tool was done. All images were cropped to include teeth and around 3 mm of surrounding tissues. Preprocessing was done using Python
- iii. Step 3: Uploading the dataset of images on Mendeley data repository

4.2. Materials






Table 2.

Table 2
Materials used to capture the images for the Dental Dataset.

Sr. No.	Material	Description	Image
1.	iPhone 15 smartphone	<ul style="list-style-type: none"> • camera with 48MP Main: 26 mm, $f/1.6$ aperture, • 100 % Focus Pixels, • super-high-resolution photos (24MP and 48MP) 	



(continued on next page)

Table 2
(continued)

Sr. No.	Material	Description	Image
2.	Cheek retractor	A plastic device that pulls the lips and cheeks away from the teeth to expose them for dental work.	
3.	Intraoral photographic mirror	Made of 18:8 stainless steel, they are used in dentistry to expose areas of a patient's mouth that are hard to see for photography.	
4.	Surface Disinfectant (70 % isopropyl alcohol)	It is a colorless, clear, bitter liquid with a sharp, musty odor that has disinfectant properties. It is extensively used as a disinfectant in hospitals, clinics and pharmacies for instruments, equipments as well as surfaces.	
5.	Containers for instruments	Instrument containers are used to transport and store medical instruments and supplies in a sterile environment.	
6.	Disposable Face masks	A single-use, three-layered mask that is part of personal protection equipment is called a surgical mask. It aids in preserving a physical barrier between the operator's mouth and nose and any possible pollutants in the surrounding area.	

(continued on next page)

Table 2
(continued)

Sr. No.	Material	Description	Image
7.	Nitrile gloves	Nitrile gloves are non latex gloves hence hypoallergenic in nature. They are considered the most protective type of disposable gloves which are puncture-proof, tear-proof, chemical-resistant, and non-irritating.	
8.	Sterile Cotton Balls	Sterile cotton balls are made from high-grade fiber that is absorbent and pure white. They are consistently sized, weighted, and colored and undergo a sterilization process to remove harmful microorganisms.	

4.3. Methods

All the images were photographed with a Smartphone after tooth cleaning and drying. To avoid condensation and ensure optimally sharp images, all occlusal surfaces were photographed indirectly using pre-warmed intraoral mirrors. For the other photographic views a cheek retractor was used [2] as shown in Fig. 4. For all photographs, distance between the smartphone and



Fig. 4. Capturing intraoral images of child patient.

Table 3
Comparative table on datasets related to teeth.

Sr. No.	Dataset Ref. No.	Repository	Number of Images	Views	Photographs/ X rays
1.	[9]	www.kaggle.com	45	Single tooth images	Photographs of extracted teeth
2.	[10]	www.kaggle.com	100	–	X rays
3.	[11]	www.kaggle.com	93	–	X rays
4.	[1] Our Published Dataset	www.data.mendeley.com	9562	8 views of teeth	Intraoral photographs

the teeth was 5–7 cm with the lens in the same horizontal plane as the teeth. To maintain exceptional image quality, photographs with out-of-focus areas or contamination were excluded. Blurred photographs were discarded to guarantee clarity and ensure optimal presentation. Furthermore, duplicated photos from identical images were omitted to avoid redundancy in the dataset, thus ensuring consistent image framing and reducing distortion. This guarantees the clarity needed for accurate analysis by algorithms [3,6,8].

Process starts with cropping of images. Images were cropped to a 1:1 aspect ratio and rotated if necessary. Batch Image Resizer, a popular tool known for its efficiency in batch image resizing was used. This allowed handling of large image collections quickly, making it ideal for research involving image-based machine learning, image analysis and data augmentation. After resizing, images were stored and numbered in sequence [4,7]. This standardized the composition and ensured that the tooth surface occupied the central area of the frame. To guarantee consistent image quality and compatibility across the dataset, all captured images were saved in JPEG format and resized to a standard resolution of 1280×550 pixels and 1024×768 pixels. This standardization process ensures the dataset works seamlessly with various machine learning applications (Table 3).

The dataset only included images of healthy teeth or surfaces without cavities. This eliminates potential bias caused by other dental conditions. Images with non-cavity defects like enamel wear; enamel hypomineralisation or restorations were excluded to focus solely on healthy tooth surfaces. Following the selection and cleaning process, the final dataset contains 9562 high-quality, anonymized clinical photographs.

These quality control measures are crucial for creating a reliable dataset that can be effectively used to train and evaluate machine learning algorithms for dental assessment. By ensuring consistency, clarity and a focus on healthy teeth, the dataset provides a strong foundation for developing accurate and unbiased AI tools for dentists [4,5].

Limitations

Expanding the dataset to encompass children from diverse geographical locations would further enhance its overall diversity and applicability.

Ethics Statement

Before commencement of the study, the Institutional Ethics Committee approved the basic premise of the study (Ethics Committee Registration No. EC/NEW/INST/2021/MH/0029 issued under New Drugs and Clinical Trials Rules, 2019). Ethical Committee Approval Letter is numbered BVDU/IEC/R1/01/23-24 dated 31st August 2023. Data collection will be performed in accordance with the Declaration of Helsinki.

Dr. Preetam Shah, Professor and PhD Guide, Department of Pedodontics and Preventive Dentistry, Bharati Vidyapeeth Deemed to be University Dental College and Hospital, cross verified each image according to the protocols to ensure optimal images in the dataset.

Since this is a noninvasive study involving clicking of photographs in cooperative children, involving only the use of minimal instruments intraorally, likelihood of any adverse event is less than minimal. Inadvertent soft tissue injury due to the instruments could occur rarely. In case of an adverse event, it will be reported as per the Department Protocol including written documentation of the date and time of the adverse event, its management, precautions to be taken henceforth for its prevention and the signature of the supervising staff.

Data Availability

[Teeth or Dental Image Dataset \(Original data\)](#) (Mendeley Data)

CRedit Author Statement

Shweta Chaudhary: Conceptualization, Writing – review & editing, Methodology, Data curation; **Preetam Shah:** Supervision; **Priyanka Paygude:** Conceptualization, Writing – review & editing, Methodology, Data curation, Supervision; **Shwetambari Chiwhane:** Writing – review & editing; **Pratibha Mahajan:** Writing – review & editing; **Prashant Chavan:** Writing – review & editing; **Manisha Kasar:** Writing – review & editing.

Acknowledgements

- This research was funded by “The Research Support Fund of Symbiosis International (Deemed University), Pune, Maharashtra, India
- We are grateful to the Principal and students of Bharati Vidyapeeth English Medium High School and the patients of Department of Pedodontics and Preventive Dentistry, Bharati Vidyapeeth Deemed to be University Dental College and Hospital, Pune for their support and cooperation in collecting the data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have potentially influenced the dataset acquisition reported in this paper.

References

- [1] D. Chaudhary, Shweta; Paygude, P.; Shah, P. (2024), “Teeth or Dental image dataset”, Mendeley Data, V1, doi: [10.17632/6zsnhrds9t.1](#).
- [2] P. Biggs, et al., Development of a methodology for the standardisation and improvement of ‘Smartphone’ photography of patterned bruises and other cutaneous injuries, *Sci. Just.* 53 (3) (2013) 358–362.
- [3] Duong et al. Automated caries detection with smartphone color photography using machine Learning, *Health Informatics Journal* 1–17 doi: [10.1177/14604582211007530](#).
- [4] F.Y. Osisanwo, J.E.T. Akinsola, O. Awodele, J.O. Hinmikaiye, O. Olakanmi, J. Akinjobi, Supervised machine learning algorithms: classification and comparison, *Int. J. Comput. Trends Technol.* 48 (3) (2017) 128–138.
- [5] Zhang, et al., Development and evaluation of deep learning for screening dental caries from oral photographs, *Oral Dis.* 28 (1) (2022) 173–181, doi:[10.1111/odi.13735](#).
- [6] S. Thite, Y. Suryawanshi, K. Patil, P. Chumchu, Coconut (Cocos nucifera) tree disease dataset: a dataset for disease detection and classification for machine learning applications, *Data Br.* 51 (2023) 109690, doi:[10.1016/j.dib.2023.109690](#).
- [7] P. Paygude, M. Gayakwad, D. Wategaonkar, R. Pawar, R. Pujeri, R.Joshi, Dried Fish dataset for Indian Seafood: A machine learning application, 2024, 110563. <https://doi.org/10.1016/j.dib.2024.110563>.

- [8] S. Kothari, S. Chiwhane, S. Jain, M. Baghel, Cancerous brain tumor detection using hybrid deep learning framework, Indonesian J. Electr. Eng. Comput. Sci. 26 (3) (2022) 1651 License CC BY-NC 4.0, doi:[10.11591/ijeecs.v26.i3.pp1651-1661](https://doi.org/10.11591/ijeecs.v26.i3.pp1651-1661).
- [9] https://www.kaggle.com/datasets/pushkar34/teeth-dataset?select=teeth_dataset.
- [10] <https://www.kaggle.com/datasets/thunderpede/panoramic-dental-dataset?select=images>.
- [11] <https://www.kaggle.com/datasets/truthisneverlinear/childrens-dental-panoramic-radiographs-dataset>.