2025 AIdentifyAGE- Dental Biological Marker

Dataset Description Models

1. Imaging Modality

Ortopantomography.

Preferred name: Digital OPG

2. Annotation Pattern

The annotation pattern followed a whole-study labelling approach. Each radiographic examination received a single ground-truth label corresponding to the individual's chronological age. These labels were subsequently used both for continuous age regression modelling and for categorical age-threshold classification.

3. Annotation methodology and structure

3.1 Method of annotation: Manual

3.2. Annotation output: Spreadsheet (alphanumeric)

4. Structure nomenclature and standards

Variable: Dental age in years

Name: Dental age

Definition: The assessment dental age in years

Units: years

5. Data use agreement/licensing

Non-commercial purpose.

6. Imaging file/structure set format

Radiographic examinations were exported in JPEG (.jpg) format following standardized preprocessing. Original imaging data were acquired in DICOM format and subsequently converted to high-resolution JPEG for model training.

7. Number of images
   Training 6870
   Validation 1472
   Test 1474
8. Demographic data of patients in the regression model[1].

- Training sample
  - Female (n = 3312)
    - Minimum: 60 months (5.02 years)

[1] The classification models use the same total sample, and the random split into training, validation, and testing sets yields subsets with similar demographic characteristics.

        - Mean: 183 months (15.28 years)
        - Standard Deviation: 65 months (5.41 years)
        - Maximum: 312 months (25.99 years)
    - Male (n = 3558)
        - Minimum: 60 months (5.01 years)
        - Mean: 173 months (14.41 years)
        - Standard Deviation: 63 months (5.27 years)
        - Maximum: 312 months (25.98 years)
    - Total (n = 6870)
        - Minimum: 60 months (5.01 years)
        - Mean: 178 months (14.83 years)
        - Standard Deviation: 64 months (5.35 years)
        - Maximum: 312 months (25.99 years)
- Validation sample
    - Female (n = 730)
        - Minimum: 60 months (5.04 years)
        - Mean: 184 months (15.29 years)
        - Standard Deviation: 64 months (5.37 years)
        - Maximum: 311 months (25.92 years)
    - Male (n = 742)
        - Minimum: 64 months (5.30 years)
        - Mean: 175 months (14.60 years)
        - Standard Deviation: 62 months (5.18 years)
        - Maximum: 311 months (25.88 years)
    - Total (n = 1472)
        - Minimum: 60 months (5.04 years)
        - Mean: 179 months (14.95 years)
        - Standard Deviation: 63 months (5.29 years)
        - Maximum: 311 months (25.92 years)
- Test sample
    - Female (n = 717)
        - Minimum: 61 months (5.07 years)
        - Mean: 183 months (15.25 years)
        - Standard Deviation: 64 months (5.36 years)
        - Maximum: 310 months (25.84 years)
    - Male (n = 757)
        - Minimum: 60 months (5.04 years)
        - Mean: 170 months (14.15 years)
        - Standard Deviation: 63 months (5.27 years)
        - Maximum: 311 months (25.93 years)
    - Total (n = 1474)
        - Minimum: 60 months (5.04 years)
        - Mean: 176 months (14.69 years)
        - Standard Deviation: 64 months (5.34 years)
        - Maximum: 311 months (25.93 years)

9. Image Characteristics
    - The resolution was normalized across all images
    - All burned-in protected health information (PHI) was removed.

10. Labeler demographics

Scope of annotation: Multi-institutional

Number of annotators: Three dentists

Expertise: Licensed dental practitioners with experience in dental radiographic interpretation.

11. Reference

AIdentifyAGE project is financed under grant 2024.07444.IACDC, within the scope of the investment “RE-C05-i08 – More Digital Science”, measure “RE-C05-i08.M04 – “Support a program of R&D projects for development and implementation of advanced cybersecurity, artificial intelligence and data science systems in public administration, as well as a scientific capacitation program”, within the scope of the agreement between the Recover Portugal Mission Structure and the Fundação para a Ciência e a Tecnologia I.P. (FCT), as an intermediary beneficiary.

Research Team

Cristiana Palmela Pereira (PI)
Ciência ID: 6A19-2961-E695
ORCID: https://orcid.org/0000-0002-9164-7189

Alexandre Francisco
Ciência ID: A012-A4D8-AEEF
ORCID: https://orcid.org/0000-0003-4852-1641

Ana Margarida Rodrigues
Ciência ID: 3616-482B-D325
ORCID: https://orcid.org/0000-0003-3081-0875

Andreas Schmeling
Ciência ID: 9315-8412-274E
ORCID ID: http://orcid.org/0000-0002-0135-1617

António Figueiras
Ciência ID: BC11-0BDD-7A8B
ORCID ID: https://orcid.org/0009-0005-6693-5137

Arlindo Oliveira
Ciência ID: E718-A5FB-4F7D
ORCID: https://orcid.org/0000-0001-8638-5594

Cátia Vaz
Ciência ID: 8718-741E-BBD9
ORCID ID: https://orcid.org/0000-0001-6074-3074

Diana Augusto
Ciência ID: A412-3A16-B200
ORCID ID: https://orcid.org/0000-0003-2698-6210

Francisco Salvado e Silva
Ciência ID: A419-CBD5-656A
ORCID ID: https://orcid.org/0000-0002-1095-8090

Henrique Oliveira
Ciência ID: D317-9198-B523
ORCID ID: https://orcid.org/0000-0001-8687-4291

Hrvoje Brkić
Ciência ID: A618-CC40-E01D
ORCID ID: https://orcid.org/0000-0003-4284-7739

José Rui Figueira
Ciência ID: 7919-0602-21AA
ORCID ID: https://orcid.org/0000-0001-9822-1167

António Miguel Areias Dias Amaral
Ciência ID: 6319-E9D4-73AE
ORCID ID: https://orcid.org/0000-0003-4629-1723

Mohamed Elbawab
Ciência ID: 9C1F-6831-A47A
ORCID ID: https://orcid.org/0000-0001-7059-6848

Paulo Guerra
Ciência ID: A313-D748-52B4
ORCID ID: https://orcid.org/0009-0002-8872-7708

Renato Marcelo
Ciência ID: 2C11-D15E-819C
ORCID ID: https://orcid.org/0009-0008-9801-2355

Rui Cardoso
Ciência ID: DE19-0E7E-F76E
ORCID ID: https://orcid.org/0009-0002-3205-7370

Rui Santos
Ciência ID: 7610-8B27-8044
ORCID ID: https://orcid.org/0000-0002-7371-363X

Valon Nushi
Ciência ID: 9E12-617C-617A
ORCID ID: https://orcid.org/0000-0002-6754-3988