Systematic Literature Review Protocol

Investigating the Employment of Machine Learning Techniques in the automated age assessment of youth

Blekinge Institute of Technology

Department of Health (TIHA)

Department of Computer Science and Engineering (DIDD)

Version 2.5

June 2019

SUMMARY

1. Background

Youth age assessment is important in a number of scenarios like: the diagnosis of endocrinological diseases, hormonal therapy follow-up, prognosis and therapies related to height prediction, criminal and civil legal matters, non-hospital births, lack of birth registration, and prevention of fraud in sports competitions [1-4].

With the number of asylum seekers growing in Europe the topic of age assessment has gained more attention, particularly considering the matter of unaccompanied minors, in which Sweden and Norway top in their reception, receiving approximately one third of the unaccompanied minors of Europe [5]. International policies for asylum seekers and refugees grant minors more protection and support, influencing heavily reception arrangements, access to healthcare and education, right to a legal guardian, right to not be detained unless in specific scenarios etc [5,6]. To guarantee the fair treatment of these unaccompanied minors seeking asylum, the establishment of the correct chronological age is imperative. The problems concerning this issue are mostly related to the lack of official identification and birth certificates [4,5]. Even when this is not the case, the authorities of the receiving country often question the validity of the documents presented to prove the chronological age [5].

In the light of this rising number of necessary age assessment procedures it becomes beneficial to make use of more automated methods. In this scenario, the machine learning (ML) techniques can be promising. ML consists of a group of prediction techniques that find patterns in data and that operate in the following way: firstly, there is a training phase in which the technique learns from a set of examples on how to perform a task (e.g. classification, clusterization), then in face of a completely new set of data this task can be correctly performed [7].

The most commonly used methods for age assessment address the evaluation of medical images taken from a body region, which are known to present chronological growth signs [6]. In this scenario, ML techniques can be promising in analyzing medical images to learn how to automatically assess these growth signs.

2. Research Protocol

The main question this systematic literature review aims to answer is: **"How the machine learning techniques are being employed in studies concerning age assessment of youth (10 to 30 years)?"**. This main question can be decomposed in the questions below:

- RQ1: Which machine learning or model building techniques are being used in the age assessment studies?
- RQ2: What data characteristics (database's origin, data collection mechanism, and ages) are being considered in the age assessment studies?
- RQ3: What type of medical imaging are being used in the studies?

- RQ4: What are the regions of interest being explored for the age assessment studies (hand, wrist, knee, etc) and what were the methods used to assess them?
- RQ5: What are the performances of the proposed methods (meta-analysis)?

2.1 Search Process

As for the search process used in this systematic review of the literature, automated searches will be run on **Pubmed**, **Web of Science** and **Scopus** databases.

The search string used in this SLR was structured using the PICO approach [8], in which the main question is decomposed in four parts: population of interest intervention, comparison and outcome. The "comparison" component was not used due as the nature of the SLR is a characterization. The components used for the automated searches are characterized below, and the change record of the complete search string with its pre-execution results can be found on the Appendix 1.

- Population: Studies involving age assessment in youth.
- Intervention: Use of medical imaging.
- Outcome: Machine learning models for age assessment

Dates for the searches: Search 1: 2018-03-21; Search 2: 2019-02-06

2.2 Study Selection – Search 1

The figure 1 presents the steps to be followed, which are adapted from the PRISMA Statement four-phase process.

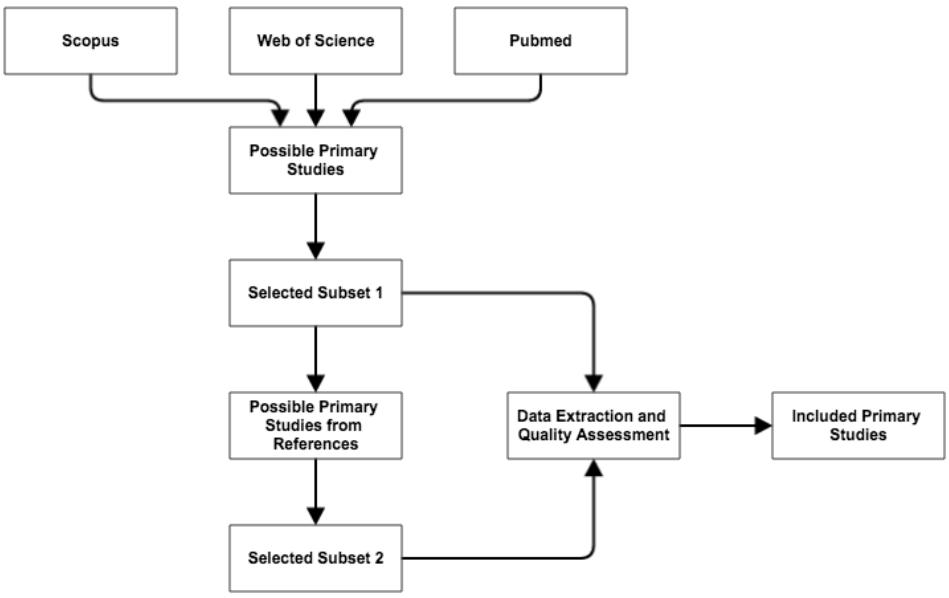

FIGURE 1. DIAGRAM FOR THE SYSTEMATIC LITERATURE REVIEW

Firstly, we will have the automated searches on the selected databases and removal of duplicates. This set of possible primary studies will be assessed by the evaluation of the titles and abstracts according to the defined inclusion and exclusion criteria.

There will be always two participants evaluating each study. This will be done in the following manner: One participant will assess all of the studies and the others will assess one third of the results, each.

The result from this evaluation is the Selected Subset 1. Then one round of backward snowballing will be carried out on the references list of the papers in the Selected Subset 1, resulting in the Selected Subset 2.

The papers in the Selected Subset 1 and 2 will be fully read and assessed for its quality using the quality assessment method specified in the section 2.3 and inclusion/exclusion criteria. The papers that successfully pass the established criteria will be part of the Included Primary Studies set and have its relevant data extracted.

Inclusion Criteria:

- Be a primary study in English; AND
- Have been published in the last 10 years; AND
- Address the research on age assessment using medical imaging; AND
- Be a study concerning the building of models for the purpose of age assessment using at least one machine learning technique; AND
- Be a study in which the age assessment is analyzed through growth zones in joints; AND
- Be a study regarding age assessment in living individuals.

Exclusion Criteria:

- Be a secondary or tertiary study; OR
- Have been published before 2007; OR
- Be written on another language other than English; OR
- Do not address research on age assessment using medical imaging; OR
- Do not address the building of models for the purpose of age assessment, using at least one machine learning technique; OR
- Be a study concerning height prediction, or a specific syndrome/disease that affects normal growth; OR
- Be a study in which the age assessment is not analyzed through growth zones in joints; OR
- Be a study in which the study population is within the range of 10 to 30 years old by less than 3 years; OR
- Be a study in which the study population is post-mortem material.

2.3 Study Selection – Search 2

The process for the selection of studies for the search 2 follows the same steps as in the search 1, without the snowballing step.

2.4 Study Quality Assessment

The quality assessment procedure checks the methodological rigor of the study and the quality of the evidence provided by it. The papers selected to be fully read will have to pass through a quality assessment in which a quality checklist will be used (table 1). Papers evaluated under 80% of the total points will not be considered to be included in the set of primary studies. The conflicts found will be resolved in a meeting including all participants.

TABLE 1. PAPER QUALITY ASSESSMENT CHECKLIST ADAPTED FROM [9] AND [10]

Design and Reporting of the Study	
Are the aims of the study clearly stated?	Yes=1 ; Partly=0.5 ; No=0
Does the study describe clearly the population being studied?	Yes=1 ; Partly=0.5 ; No=0
Does the study describe clearly the inclusion/exclusion criteria?	Yes=1 ; Partly=0.5 ; No=0
Are the limitations of the study reported either during the explanation of the study design or during the discussion of the study results?	Yes=1 ; Partly=0.5 ; No=0
Were the findings clearly reported?	Yes=1 ; Partly=0.5 ; No=0
Is there no presence of a lack of blinding that could introduce bias?	Yes=1 ; Partly=0.5 ; No=0
Data Quality	
Are the data characteristics (variables, subjects' country of origin etc) used in the study clearly defined?	Yes=1 ; Partly=0.5 ; No=0
Are the data characteristics (variables, subjects' country of origin etc) used in the study valid?	Yes=1 ; Partly=0.5 ; No=0
Is the data collection method clearly described?	Yes=1 ; Partly=0.5 ; No=0
Are the data characteristics related to age valid/verified in some way?	Yes=1 ; Partly=0.5 ; No=0
Technique(s) Employed	
Is/are the ML technique(s) being employed clearly described?	Yes=1 ; Partly=0.5 ; No=0
Is it clear how accuracy was measured?	Yes=1 ; Partly=0.5 ; No=0
Was statistical significance assessed?	Yes=1 ; Partly=0.5 ; No=0
TOTAL	**/ 13**

2.5 Data Extraction

The papers that successfully passed the inclusion/exclusion criteria and quality assessment will have their data extracted as in the tables below.

TABLE 2. DATA EXTRACTION FORM

ID			Extracted by	
Title			Journal/Source	
Authors			Year	
Publication Type				
Aim of the Study				
Dataset Size				
Age Range Male			Age Range Female	
Country of origin of the subjects of the study				
Employed machine learning techniques and other relevant statistical analyses				
Data characteristics				
Type of image (ex: x-ray, MRI etc)				
Regions of interest (hand, wrist etc)				
Methods used for the age assessment (ex: staging, comparison to reference images, computer vision)				
Dataset size				
Best achieved performance				

2.6 Data Synthesis, analysis and Reporting

The data synthesis phase will compile the data extracted from the selected studies in summary tables in order to answer the posed questions and to facilitate the following reporting process. The meta-analysis of the bone age assessment performances was done through the average of the performances weighted by the sample sizes. The software used to perform the data analyses was Excel.

REFERENCES

[1] Ekizoglu, Oguzhan, et al. "Magnetic resonance imaging of distal tibia and calcaneus for forensic age estimation in living individuals." International journal of legal medicine 129.4 (2015): 825-831.

[2] Ekizoglu, Oguzhan, et al. "Forensic age estimation via 3-T magnetic resonance imaging of ossification of the proximal tibial and distal femoral epiphyses: Use of a T2-weighted fast spin-echo technique." Forensic science international 260 (2016): 102-e1.

[3] Terada, Yasuhiko, et al. "Skeletal age assessment in children using an open compact MRI system." Magnetic resonance in medicine 69.6 (2013): 1697-1702.

[4] Urschler, Martin, Sabine Grassegger, and Darko Štern. "What automated age estimation of hand and wrist MRI data tells us about skeletal maturation in male adolescents." Annals of human biology 42.4 (2015): 358-367.

[5] Hjern, Anders, Maria Brendler-Lindqvist, and Marie Norredam. "Age assessment of young asylum seekers." Acta paediatrica 101.1 (2012): 4-7.

[6] UNHCR. "Guidelines on policies and procedures in dealing with unaccompanied children seeking asylum". Geneva: UNHCR 1997.

[7] Louridas P., Ebert C. Machine Learning. IEEE Softw. 2016 Sep; 33(5):110–5.

[8] Pai, Madhukar, et al. "Systematic reviews and meta-analyses: an illustrated, step-by-step guide." *The National medical journal of India* 17.2 (2003): 86-95.

[9] B. Kitchenham. "*Guidelines for performing systematic literature reviews in software engineering (version 2.3)"*. Software Engineering Group, School of Computer Science and Mathematics, Keele University and Department of Computer Science, University of Durham, July 2007.

[10] SBU - Statens Beredning för Medicinsk och Social Utvärdering. "Mall för bedömning av relevans". 2014. Available at: http://www.sbu.se/globalassets/ebm/metodbok/mall_relevans.pdf. Last accessed on 2017-12-08.

APPENDIX 1 – CHANGE RECORD OF THE PICO SEARCH STRING

String 1

- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
143	534	44

((“age assessment” OR “skeletal maturity” OR “bone age assessment” OR “bone maturity” OR “bone development” OR “bone age testing” OR “bone age tests” OR “chronological age estimations” OR “skeletal maturation” OR “dental maturity” OR “dental age assessment” OR “dental examination” OR “bone examination” OR “skeletal development” OR “age testing” OR “physical development” OR “developmental assessments” OR “age diagnostics” OR “bone-age” OR “skeletal age”) **AND (**“Magnetic Resonance imaging” OR “MRI”**)**

AND

(”Classification” OR ”Regression” OR ”Kernel” OR ”Support vector machines” OR ”Gaussian process” OR ”Neural networks” OR ”Logical learning” OR ”relational learning” OR ”Inductive logic” OR ”Statistical relational” OR ”probabilistic graphical model” OR ”Maximum likelihood” OR ”Maximum entropy” OR ”Maximum a posteriori” OR ”Mixture model” OR ”Latent variable model” OR ”Bayesian network” OR ”linear model” OR ”Perceptron algorithm” OR ”Factorization” OR ”Factor analysis” OR ”Principal component analysis” OR ”Canonical correlation” OR ”Latent Dirichlet allocation” OR ”Rule learning” OR ”Instance-based” OR ”Markov” OR ”Stochastic game” OR ”Learning latent representation” OR ”Deep belief network” OR ”Bio-inspired approach” OR ”Artificial life” OR ”Evolvable hardware” OR ”Genetic algorithm” OR ”Genetic programming” OR ”Evolutionary robotic” OR ”Generative and developmental approaches”))

String 2

- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
8216	5256	18769

(
(((“age” OR “bone age” OR “skeletal” OR “dental”) AND (“assessment” OR “appraisal” OR “estimate” OR “estimation” OR “determination” OR “prediction” OR “test” OR “testing”)) OR “age diagnostics”)
AND
(“Magnetic Resonance imaging” OR “MRI”**)**
AND
(”Classification” OR ”Regression” OR ”Kernel” OR ”Support vector machines” OR ”Gaussian process” OR ”Neural networks” OR ”Logical learning” OR ”relational learning” OR ”Inductive logic” OR ”Statistical relational” OR ”probabilistic graphical model” OR ”Maximum likelihood” OR ”Maximum entropy” OR ”Maximum a posteriori” OR ”Mixture model” OR ”Latent variable model” OR ”Bayesian network” OR ”linear model” OR ”Perceptron algorithm” OR ”Factorization” OR ”Factor analysis” OR ”Principal component analysis” OR ”Canonical correlation” OR ”Latent Dirichlet allocation” OR ”Rule learning” OR ”Instance-based” OR ”Markov” OR ”Stochastic game” OR ”Learning latent representation” OR ”Deep belief network” OR ”Bio-inspired approach” OR ”Artificial life” OR ”Evolvable hardware” OR ”Genetic algorithm” OR ”Genetic programming” OR ”Evolutionary robotic” OR ”Generative and developmental approaches”)
)

String 3

- **Notes:** Added terms in red, remove “physical development”
- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
22	43	55

((“age assessment” OR “age appraisal” OR “age diagnostics” OR “age estimate” OR “age estimation” OR “age determination” OR “age prediction” OR “age testing” OR “bone age assessment” OR “bone maturity” OR “bone development” OR “bone age testing” OR “bone age tests” OR “skeletal maturity” “skeletal maturation” OR “dental maturity” OR “dental age assessment” OR “dental examination” OR “bone examination” OR “skeletal development” OR “developmental assessment” OR “bone-age” OR “skeletal age”)
AND
(“Magnetic Resonance imaging” OR “MRI”**)**
AND
(”Classification” OR ”Regression” OR ”Kernel” OR ”Support vector machines” OR ”Gaussian process” OR ”Neural networks” OR ”Logical learning” OR ”relational learning” OR ”Inductive logic” OR ”Statistical relational” OR ”probabilistic graphical model” OR ”Maximum likelihood” OR ”Maximum entropy” OR ”Maximum a posteriori” OR ”Mixture model” OR ”Latent variable model” OR ”Bayesian network” OR ”linear model” OR ”Perceptron algorithm” OR ”Factorization” OR ”Factor analysis” OR ”Principal component analysis” OR ”Canonical correlation” OR ”Latent Dirichlet allocation” OR ”Rule learning” OR ”Instance-based” OR ”Markov” OR ”Stochastic game” OR ”Learning latent representation” OR ”Deep belief network” OR ”Bio-inspired approach” OR ”Artificial life” OR ”Evolvable hardware” OR ”Genetic algorithm” OR ”Genetic programming” OR ”Evolutionary robotic” OR ”Generative and developmental approaches”))

String 4

- **Notes:** Added terms in red
- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
158	350	322

((“age assessment” OR “age appraisal” OR “age diagnostics” OR “age estimate” OR “age estimation” OR “age determination” OR “age prediction” OR “age testing” OR “bone age assessment” OR “bone maturity” OR “bone development” OR “bone age testing” OR “bone age tests” OR “skeletal maturity” “skeletal maturation” OR “dental maturity” OR “dental age assessment” OR “dental examination” OR “bone examination” OR “skeletal development” OR “developmental assessment” OR “bone-age” OR “skeletal age”)
AND
(“Magnetic Resonance imaging” OR “MRI” OR “x ray” OR “x-ray” OR “xray” OR “computed tomography” OR “CT” OR “ultrasound”**)**
AND
(”Classification” OR ”Regression” OR ”Kernel” OR ”Support vector machines” OR ”Gaussian process” OR ”Neural networks” OR ”Logical learning” OR ”relational learning” OR ”Inductive logic” OR ”Statistical relational” OR ”probabilistic graphical model” OR ”Maximum likelihood” OR ”Maximum entropy” OR ”Maximum a posteriori” OR ”Mixture model” OR ”Latent variable model” OR ”Bayesian network” OR ”linear model” OR ”Perceptron algorithm” OR ”Factorization” OR ”Factor analysis” OR ”Principal component analysis” OR ”Canonical correlation” OR ”Latent Dirichlet allocation” OR ”Rule learning” OR ”Instance-based” OR ”Markov” OR ”Stochastic game” OR ”Learning latent representation” OR ”Deep belief network” OR ”Bio-inspired approach” OR ”Artificial life” OR ”Evolvable hardware” OR ”Genetic algorithm” OR ”Genetic programming” OR ”Evolutionary robotic” OR ”Generative and developmental approaches”))

String 5

- **Notes:** Added terms in red
- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
168	509	324

((“age assessment” OR “age appraisal” OR “age diagnostics” OR “age estimate” OR “age estimation” OR “age determination” OR “age prediction” OR “age testing” OR “bone age assessment” OR “bone maturity” OR “bone development” OR “bone age testing” OR “bone age tests” OR “skeletal maturity” “skeletal maturation” OR “dental maturity” OR “dental age assessment” OR “dental examination” OR “bone examination” OR “skeletal development” OR “developmental assessment” OR “bone-age” OR “skeletal age” OR “growth zone”)
AND
(“Magnetic Resonance imaging” OR “MRI” OR “x ray” OR “x-ray” OR “xray” OR “computed tomography” OR “CT” OR “ultrasound” OR “medical imaging”**)**
AND
(”Classification” OR ”Regression” OR ”Kernel” OR ”Support vector machines” OR ”Gaussian process” OR ”Neural networks” OR ”Logical learning” OR ”relational learning” OR ”Inductive logic” OR ”Statistical relational” OR ”probabilistic graphical model” OR ”Maximum likelihood” OR ”Maximum entropy” OR ”Maximum a posteriori” OR ”Mixture model” OR ”Latent variable model” OR ”Bayesian network” OR ”linear model” OR ”Perceptron algorithm” OR ”Factorization” OR ”Factor analysis” OR ”Principal component analysis” OR ”Canonical correlation” OR ”Latent Dirichlet allocation” OR ”Rule learning” OR ”Instance-based” OR ”Markov” OR ”Stochastic game” OR ”Learning latent representation” OR ”Deep belief network” OR ”Bio-inspired approach” OR ”Artificial life” OR ”Evolvable hardware” OR ”Genetic algorithm” OR ”Genetic programming” OR ”Evolutionary robotic” OR ”Generative and developmental approaches”))

String 6

- **Notes:** Added terms in red
- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
329	528	369

((“age assessment” OR “age appraisal” OR “age diagnostics” OR “age estimate” OR “age estimation” OR “age determination” OR “age prediction” OR “age testing” OR “bone age measurement” OR “bone age assessment” OR “bone maturity” OR “bone development” OR “bone age testing” OR “bone age tests” OR “skeletal maturity” “skeletal maturation” OR “dental maturity” OR “dental age assessment” OR “dental examination” OR “bone examination” OR “skeletal development” OR “developmental assessment” OR “bone age” OR “skeletal age” OR “growth zone”) **AND (**“Magnetic Resonance imaging” OR “MRI” OR “x ray” OR “x-ray” OR “xray” OR “Radiography” OR “computed tomography” OR “CT” OR “ultrasound” OR “ultrasonography” OR “medical imaging”**) AND** (“machine learning” OR “unsupervised Machine Learning” OR “supervised Machine Learning” OR ”Classification” OR ”Regression” OR ”Kernel” OR ”Support vector machines” OR ”Gaussian process” OR ”Neural networks” OR ”Logical learning” OR ”relational learning” OR ”Inductive logic” OR ”Statistical relational” OR ”probabilistic graphical model” OR ”Maximum likelihood” OR ”Maximum entropy” OR ”Maximum a posteriori” OR ”Mixture model” OR ”Latent variable model” OR ”Bayesian network” OR ”linear model” OR ”Perceptron algorithm” OR ”Factorization” OR ”Factor analysis” OR ”Principal component analysis” OR ”Canonical correlation” OR ”Latent Dirichlet allocation” OR ”Rule learning” OR ”Instance-based” OR ”Markov” OR ”Stochastic game” OR ”Learning latent representation” OR ”Deep belief network” OR ”Bio-inspired approach” OR ”Artificial life” OR ”Evolvable hardware” OR ”Genetic algorithm” OR ”Genetic programming” OR ”Evolutionary robotic” OR ”Generative and developmental approaches”))

String 7

- **Notes:** Filter for the last 10 years, English papers
- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
226	274	287

((“age assessment” OR “age appraisal” OR “age diagnostics” OR “age estimate” OR “age estimation” OR “age determination” OR “age prediction” OR “age testing” OR “bone age measurement” OR “bone age assessment” OR “bone maturity” OR “bone development” OR “bone age testing” OR “bone age tests” OR “skeletal maturity” “skeletal maturation” OR “dental maturity” OR “dental age assessment” OR “dental examination” OR “bone examination” OR “skeletal development” OR “developmental assessment” OR “bone age” OR “skeletal age” OR “growth zone”)

AND
(“Magnetic Resonance imaging” OR “MRI” OR “x ray” OR “x-ray” OR “xray” OR “Radiography” OR “computed tomography” OR “CT” OR “ultrasound” OR “ultrasonography” OR “medical imaging”**) AND** (“machine learning” OR “unsupervised Machine Learning” OR “supervised Machine Learning” OR ”Classification” OR ”Regression” OR ”Kernel” OR ”Support vector machines” OR ”Gaussian process” OR ”Neural networks” OR ”Logical learning” OR ”relational learning” OR ”Inductive logic” OR ”Statistical relational” OR ”probabilistic graphical model” OR ”Maximum likelihood” OR ”Maximum entropy” OR ”Maximum a posteriori” OR ”Mixture model” OR ”Latent variable model” OR ”Bayesian network” OR ”linear model” OR ”Perceptron algorithm” OR ”Factorization” OR ”Factor analysis” OR ”Principal component analysis” OR ”Canonical correlation” OR ”Latent Dirichlet allocation” OR ”Rule learning” OR ”Instance-based” OR ”Markov” OR ”Stochastic game” OR ”Learning latent representation” OR ”Deep belief network” OR ”Bio-inspired approach” OR ”Artificial life” OR ”Evolvable hardware” OR ”Genetic algorithm” OR ”Genetic programming” OR ”Evolutionary robotic” OR ”Generative and developmental approaches”))

String 8

- **Notes:** Changes in red.
- **Pre Execution Results:**

Scopus	Pubmed	Web of Science
90	67	49

((("age assessment" OR "age appraisal" OR "age diagnostics" OR "age estimate" OR "age estimation" OR "age determination" OR "age prediction" OR "age testing") AND ("bone age measurement" OR "bone age assessment" OR "bone maturity" OR "bone development" OR "bone age testing" OR "bone age tests" OR "skeletal maturity" "skeletal maturation" OR "bone examination" OR "skeletal development" OR "developmental assessment" OR "bone age" OR "skeletal age" OR "growth zone")) **AND (**"Magnetic Resonance imaging" OR "MRI" OR "x ray" OR "x-ray" OR "xray" OR "Radiography" OR "computed tomography" OR "CT" OR "ultrasound" OR "ultrasonography" OR "medical imaging"**) AND** ("machine learning" OR "unsupervised Machine Learning" OR "supervised Machine Learning" OR "Classification" OR "Regression" OR "Kernel" OR "Support vector machines" OR "Gaussian process" OR "Neural networks" OR "Logical learning" OR "relational learning" OR "Inductive logic" OR "Statistical relational" OR "probabilistic graphical model" OR "Maximum likelihood" OR "Maximum entropy" OR "Maximum a posteriori" OR "Mixture model" OR "Latent variable model" OR "Bayesian network" OR "linear model" OR "Perceptron algorithm" OR "Factorization" OR "Factor analysis" OR "Principal component analysis" OR "Canonical correlation" OR "Latent Dirichlet allocation" OR "Rule learning" OR "Instance-based" OR "Markov" OR "Stochastic game" OR "Learning latent representation" OR "Deep belief network" OR "Bio-inspired approach" OR "Artificial life" OR "Evolvable hardware" OR "Genetic algorithm" OR "Genetic programming" OR "Evolutionary robotic" OR "Generative and developmental approaches"))