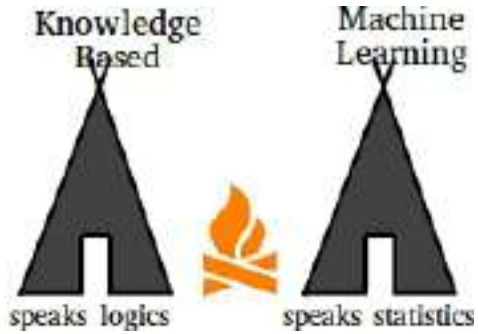
Interleaving machine learning with reasoning for identifying retinal conditions

Adrian Groza

Intelligent Systems Group, Department of Computer Science
Technical University of Cluj-Napoca, Romania

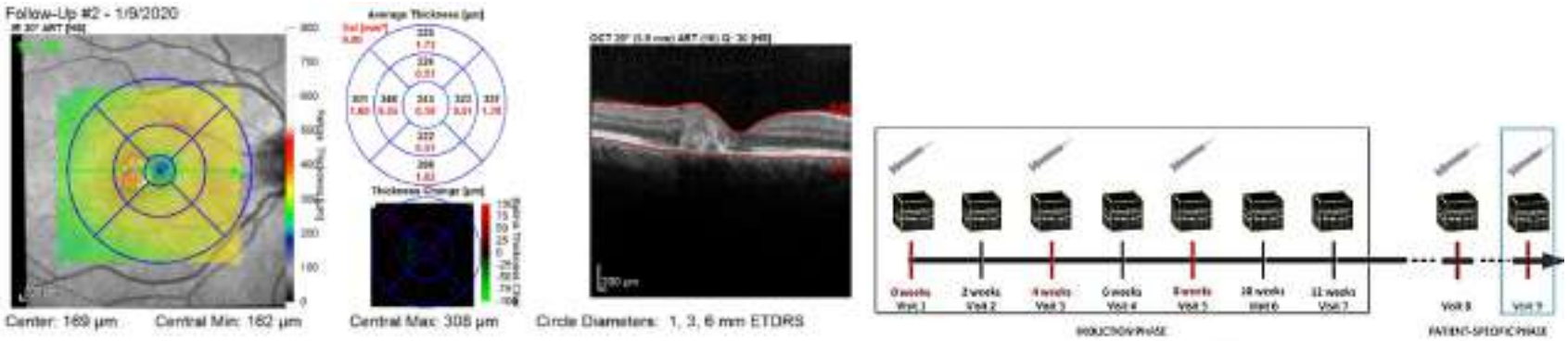


Running project (2022-2024)

New OCT Biomarkers Identified with Deep Learning for Risk Stratification of Patients with Age-related Macular Degeneration, PED616, University of Medicine an Pharmacy Iuliu Hatieganu, Cluj-Napoca (Prof. Simona Delia Nicoara)

Adrian.Groza@cs.utcluj.ro
https://isg.utcluj.ro

Predicting visual acuity from small-time series



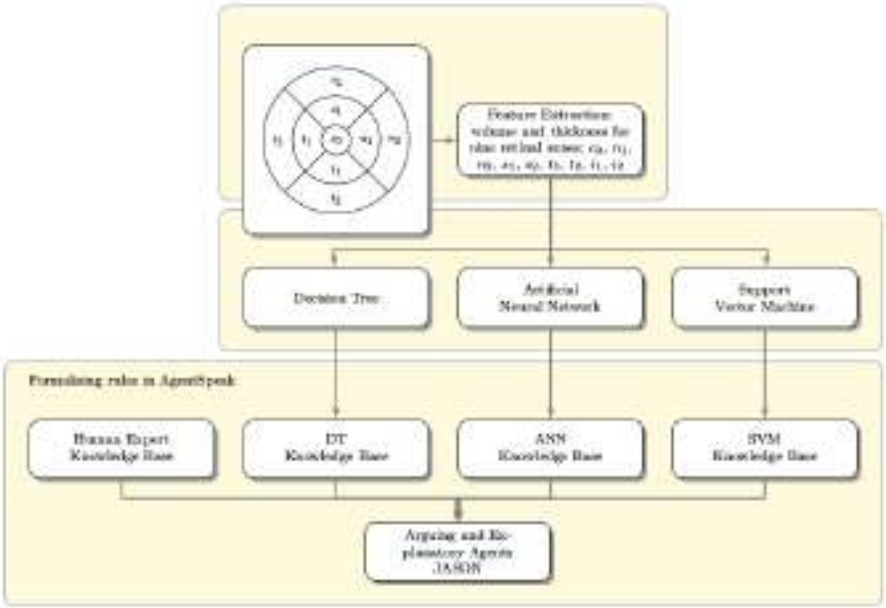
How to learn from small-sized time series? How to handle different time intervals between visits? How to learn from different numbers of visits (1–5)?
Technologies used: linear regression, gradient boosting, random forest and extremely randomised trees, bidirectional RNN, LSTM network, GRU network
Conducted Experiments considering:

1. Only previous measured visual acuity
2. Numerical OCT features, e.g. thickness and volume values in all retinal zones
3. Fundus scan images represented as embeddings obtained from the convolutional autoencoder (increased accuracy for all algorithms)

Main result: $R^2 = 0.99$, LSTM, 3 visits (monthly resampled series) based on numerical OCT values, fundus images, and previous visual acuities.

Argue on Classifications of Retinal Conditions

- explain algorithmic decisions to humans (e.g. by extracting rules from models)
- include the ophthalmologist in the loop (by including expert knowledge)
- build safety cases (by creating assurance argument patterns in Goal Structering Notation)

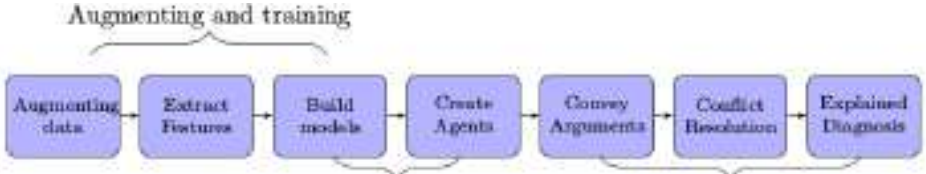


$R_1^{DT(a=.97)}$: $t(s_1) \leq .35 \wedge v(s_1) \leq .51 \rightarrow^{69} \langle 1, 0, 0 \rangle$

$R_2^{SVM(a=.7)}$: $t(n_2) \leq .45 \wedge t(t_2) > .41 \wedge v(n_2) < 2.41 \wedge v($
$1.94) \rightarrow \langle .0149, .5373, .4478 \rangle$

$R_1^{ANN(a=.75)}$: $v(t_2) \leq 1.28 \rightarrow \langle .0045, .0856, .9099 \rangle$

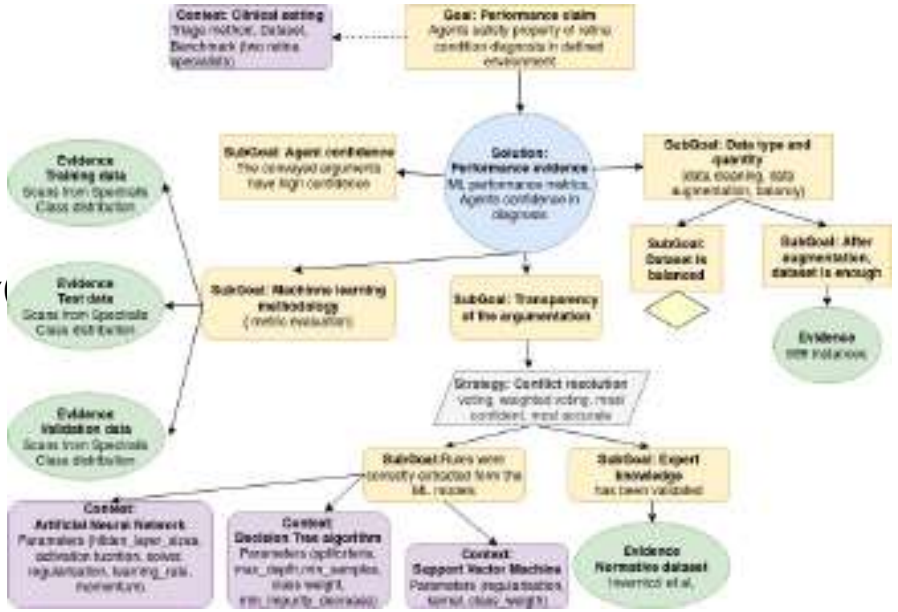
R_1^{E}: $t(c_0) = 280.1 \pm 17.5 \rightarrow^{200} \langle 0, 0, 1 \rangle$



A support tool for ophtalmologist:
Generating explanations in NL

[Master] Suggested diagnosis is *Diabetic Retinopathy*
Agent *DT* is 100% sure, and *DT*'s accuracy is 0.96
Agent *SVM* is 97% sure, and *SVM*'s accuracy is 0.75
Agent *ANN* is 95.79% sure, and *ANN*'s accuracy is 0.95
Agent *E* had no arguments

[Master] Diagnosis *Diabetic Retinopathy* was chosen because:
The thickness value in t_1 zone is smaller than 0.34 and
The thickness value in t_2 zone is smaller than 0.3 and
The thickness value in s_2 zone is greater than 0.3 and
The volume value in s_1 zone is smaller than 0.58 and
The thickness value in s_1 zone is greater than 0.35.

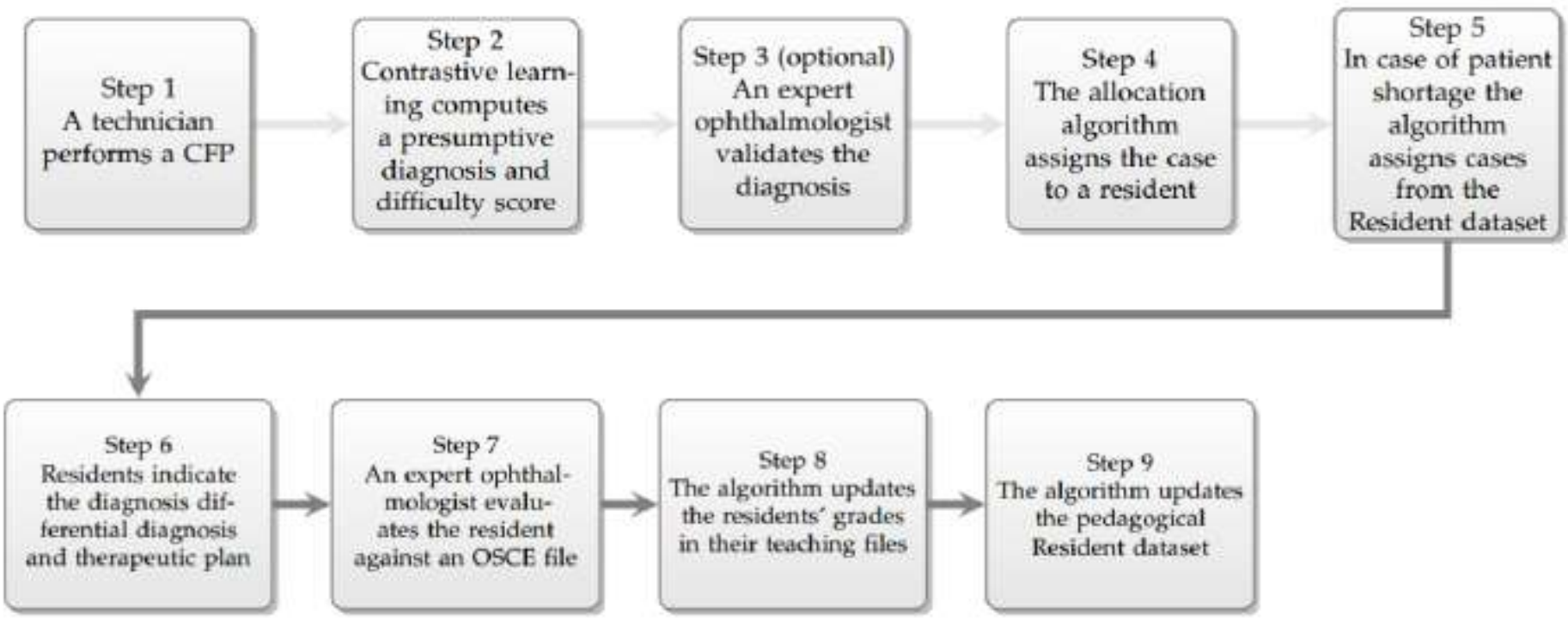


AI for personalized residency training

Given a case (a retinal condition), which resident would benefit the most?

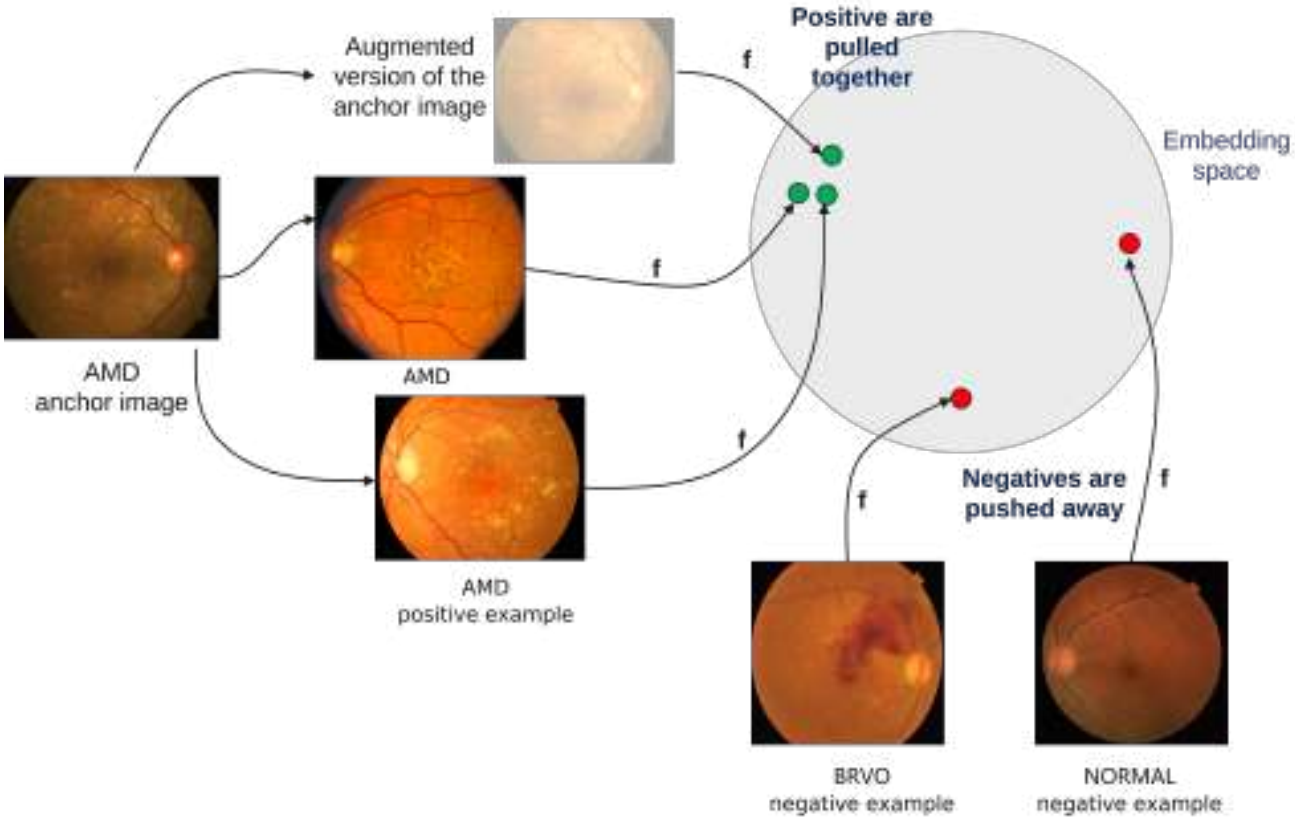
A neuro-symbolic case allocation algorithm

1. Deep learning gives a presumptive diagnosis and assesses case difficulty
2. Expert sytems allocates cases to residents



Classifying Color Fundus Photographs with Deep Learning

1. Building the Resident Training Dataset- 9.693 fundus, 19 conditions
2. Applying constrastive learning: conditions that have similar aspects are closer (e.g. drusen closer to AMD than to glaucoma)
3. Automatically assessing difficult cases



Prob difficulty (PD): how confident the model is about the predicted classes and to what extent the signs of other classes are identified without enough confidence to predict that class.
Neighbors difficulty (ND): Let $x : C$. If all x neighbors belong to $C \rightarrow x$ is easy. If none of x's the neighbors belong to C and they are very close$\rightarrow$ x is difficult. If mean $\delta(x$, neighbors from the same class$) < \delta(x$, neighbors from other classes$) \rightarrow x$ is rather easy (even though there is a diversity among the neighbors, the dominant ones are from C). Otherwise, the case is rather difficult.

Allocation rules based on expert systems (19 rules)

Educational topics in the retina study module

#	Educational Topic	Retinal Condition
T_1	Normal	Normal, tessellated fundus
T_2	Macular conditions	AMD, pathological myopia, drusen, epiretinal membrane, CSC
T_3	Vascular conditions	DR, hypertensive retinopathy, branch retinal vein occlusion, central RVO
T_4	Optic nerve conditions	Glaucoma, large optic cup, optic disc edema, myelinated nerve fibers
T_5	Peripheral retina conditions	Rhegmatogenous retinal detachment, laser spots
T_6	Transparent media conditions	Vitreous degeneration, refractive media opacity

#	Rule
r_1	Assign at least one case/day to each resident
r_2	Assign with priority patients presenting to the retina clinic, then, in case of shortage, CFPs from the RT dataset
r_3	Assign one case from each of the 19 retinal conditions to each resident
r_4	Assign the case to the resident which has seen fewer cases from this retinal conditions, up to 3 cases
r_5	Assign the case to the resident with the lowest grade (performance score + difficulty score) until grade ≥ 7
r_6	Assign the case to the resident with the oldest encounter for that specific condition
r_7	Assign the case to the resident with the lowest number of cases from that specific educational topic
r_8	Assign the case to the resident with the lowest number of cases from that specific retinal condition
r_9	Assign the case to the resident with the lowest number of cases from all the 19 retinal conditions

Intermediate spaced repetition: aiming at knowledge revision when someone is just about to forget.

Resident:		*Date:*	
	OSCE DR		
Correct Diagnosis Pass (Calculate Score)	☑	**Wrong Diagnosis Fail (0 Points)**	☐
Clinical fundus signs	*(each box = 1 point)*		
Microaneurysms	☑	Neovascularisation of the disc	☑
Dot-blot hemorrhages	☑	Neovascularisation elsewhere	☐
Hard exudates	☑	Preretinal hemorrhage	☐
Cotton-wool spots	☑	Vitreous hemorrhage	☐
Venous beading	☑	Tractional retinal detachment	☐
Intraretinal microvascular anomalies	☑	Laser spots	☐
Differential diagnosis of macular edema	*(each box = 1 point)*		
Hypertensive retinopathy	☑	Macular edema secondary to epiretinal membrane	☐
Central retinal vein occlusion	☐	Ruptured microaneurysm	☐
Branch retinal vein occlusion	☑	Irvine gass syndrome	☑
Choroidal neovascular membrane	☐	Post uveitic macular edema	☑
Differential diagnosis of retinopathy	*(each box = 1 point)*		
Central retinal vein occlusion	☑	Valsalva retinopathy	☐
Hemiretinal vein occlusion	☑	Sickle cell retinopathy	☐
Branch retinal vein occlusion	☑	Post-traumatic retinal bleed	☑
Hypertensive retinopathy	☑	Retinal macroaneurysm	☐
Ocular ischemic syndrome	☑	Retinopathy in thalassemia	☐
Terson syndrome	☑		
Management of macular edema	*(each box = 1 point)*		
Observation	☐	Intravitreal anti-VEGF	☑
Management of retinopathy	*(each box = 1 point)*		
Observation	☐	Intravitreal anti-VEGF	☑
Panfundus laser photocoagulation	☑	Vitrectomy	☐
Resident scored (29) points of a total of 37			
Physician:		*Score:* (4)	

Since the neuro-symbolic case allocation affect students learning, the system should comply with the AI Act

Reasoning on ontologies for AMD diagnosis

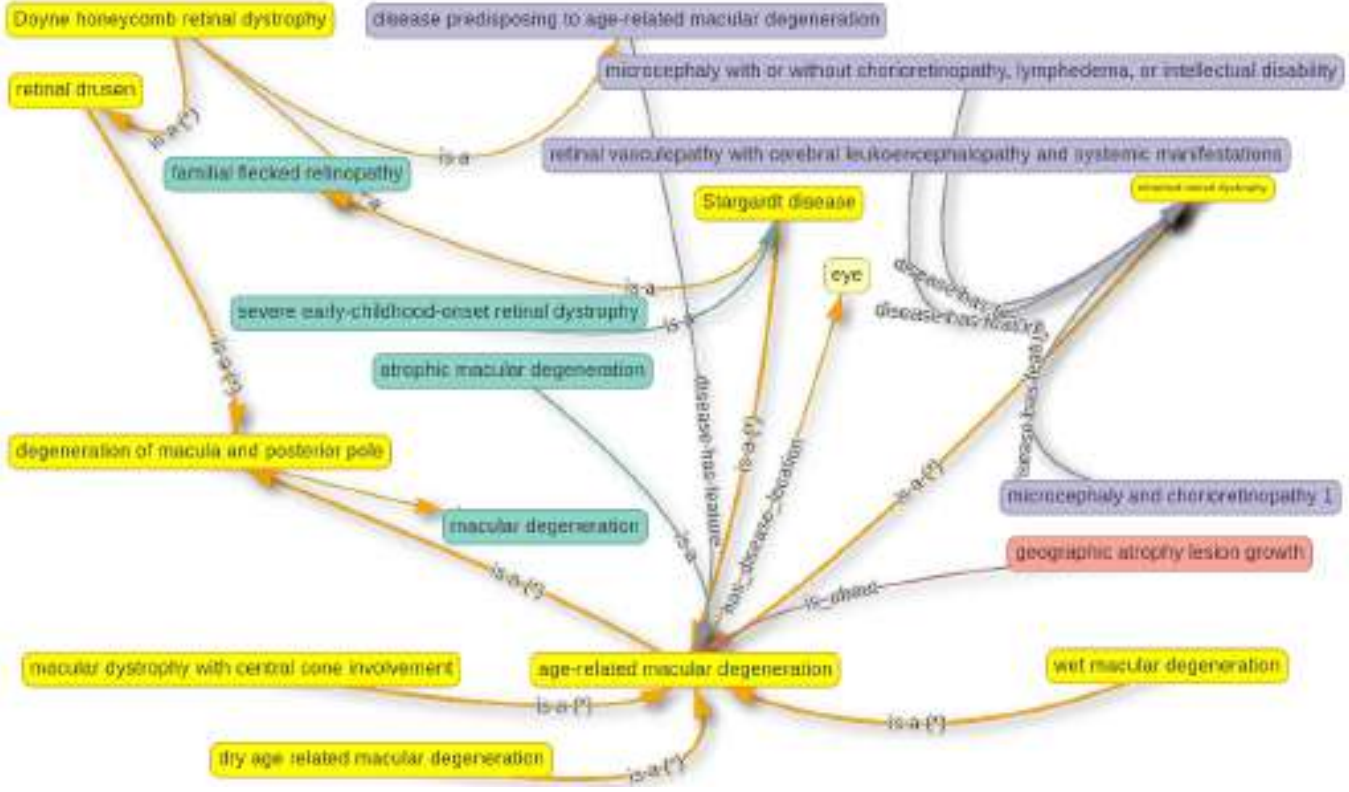
Formalising medical protocols in Description Logics

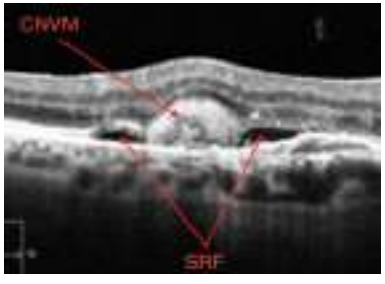
Sample of definitions and classifications scales for AMD

		Epidemiological classification (Wisconsin grading)
$EarlyAMD^W$	$\equiv$	$AMD \sqcap \exists hasBiomarker.(LargeDrusen \sqcup RetinalPseudodrusen \sqcup PigmentaryAbn)$
$LateAMD^W$	$\equiv$	$NeovascularAMD \sqcup GeographicAtropy$
		Basic clinical classification
$NoAgeingChanges^C$	$\equiv$	$\forall hasDrusen.\bot \sqcap \forall hasAbn.\neg PigmentaryAbn$
$NormalAgeingChanges^C$	$\equiv$	$\forall hasDrusen.SmallDrusen \sqcap \forall hasAbn.\neg PigmentaryAb$
$EarlyAMD^C$	$\equiv$	$AMD \sqcap \exists hasBiomarker.MediumDrusen \sqcap \forall hasAbnormalities.\neg PigmentaryAbnormalities$
$IntermediateAMD^C$	$\equiv$	$AMD \sqcap (\exists hasBiomarker.LargeDrusen \sqcap \exists hasAbnormalities.\neg PigmentaryAbnormalities$
$LateAMD^C$	$\equiv$	$NeovascularAMD \sqcup GeographicAtropyy$
		AREDS simplified severity scale points
$Severity_0$	$\equiv$	$\forall hasBiomarker.\neg LargeDrusen \sqcup \forall changes.\neg Pigment$
$Severity_1$	$\equiv$	$\exists hasBiomarker.\neg LargeDrusen \sqcup (= 1)changes.Pigment$
$Severity_4$	$\equiv$	$(> 1)hasBiomarker.LargeDrusen \sqcap (> 1)changes.Pigment \sqcap$

Describing OCT biomarkers in Description Logics

Anatomy ontology, Human Disease, Experimental Factor Ontology, SNOMED, BiologicalSpatial Ontology, Relation ontology, Symptom ontology





$$f_3 : hasDisease.WetAMD \quad (1)$$

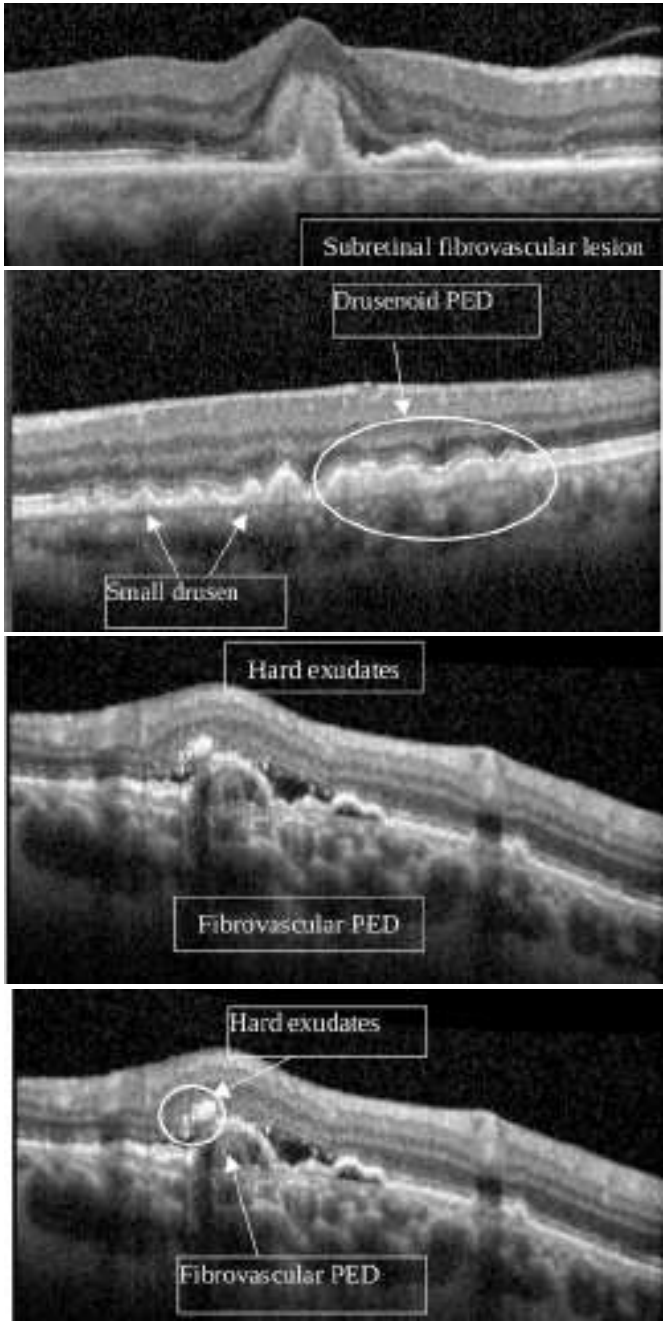
$$f_3 : \exists hasBM.(Type2CNVM \sqcap isAbove.RPE \sqcap \exists hasAdjacentBM.SRF \sqcap \forall hasAdjacentBM.SRF \quad (2)$$

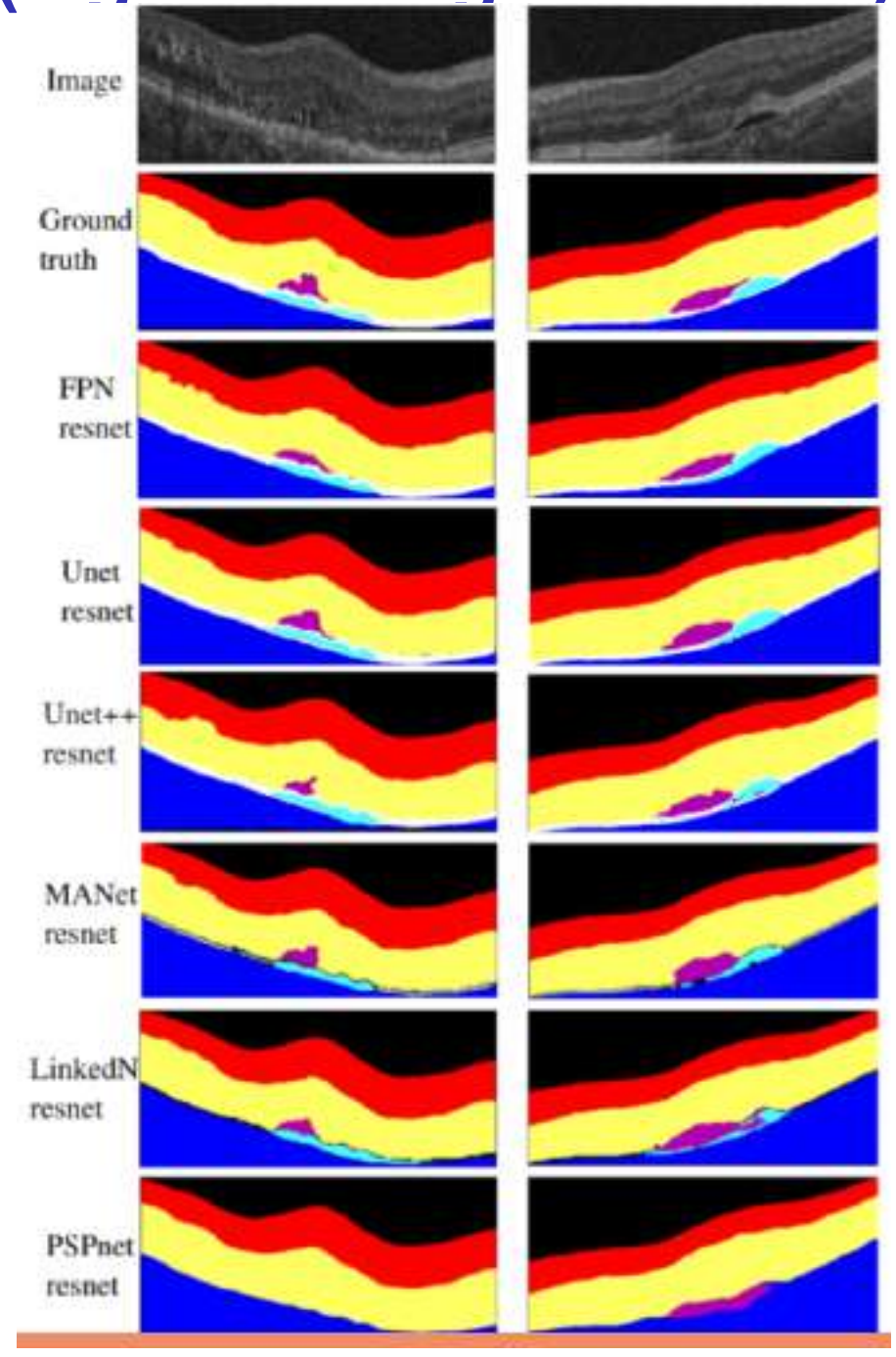
$$f_3 : \exists hasBM.(Exudate \sqcap isLocated.Nasal) \quad (3)$$

$$Type1CNVM \sqsubseteq CNVM \sqcap \exists isBeneath.RPE \sqcap$$
$$\exists appear.(Fibrovascular \sqcup HemorrhagicPigmentEpithelialDetachment) \quad (4)$$

$$Type2CNVM \sqsubseteq CNVM \sqcap \exists isAbove.RPE \exists hasAdjacentBM.SRF \quad (5)$$

Biomarker segmentation (e.g. fluid segmentation)



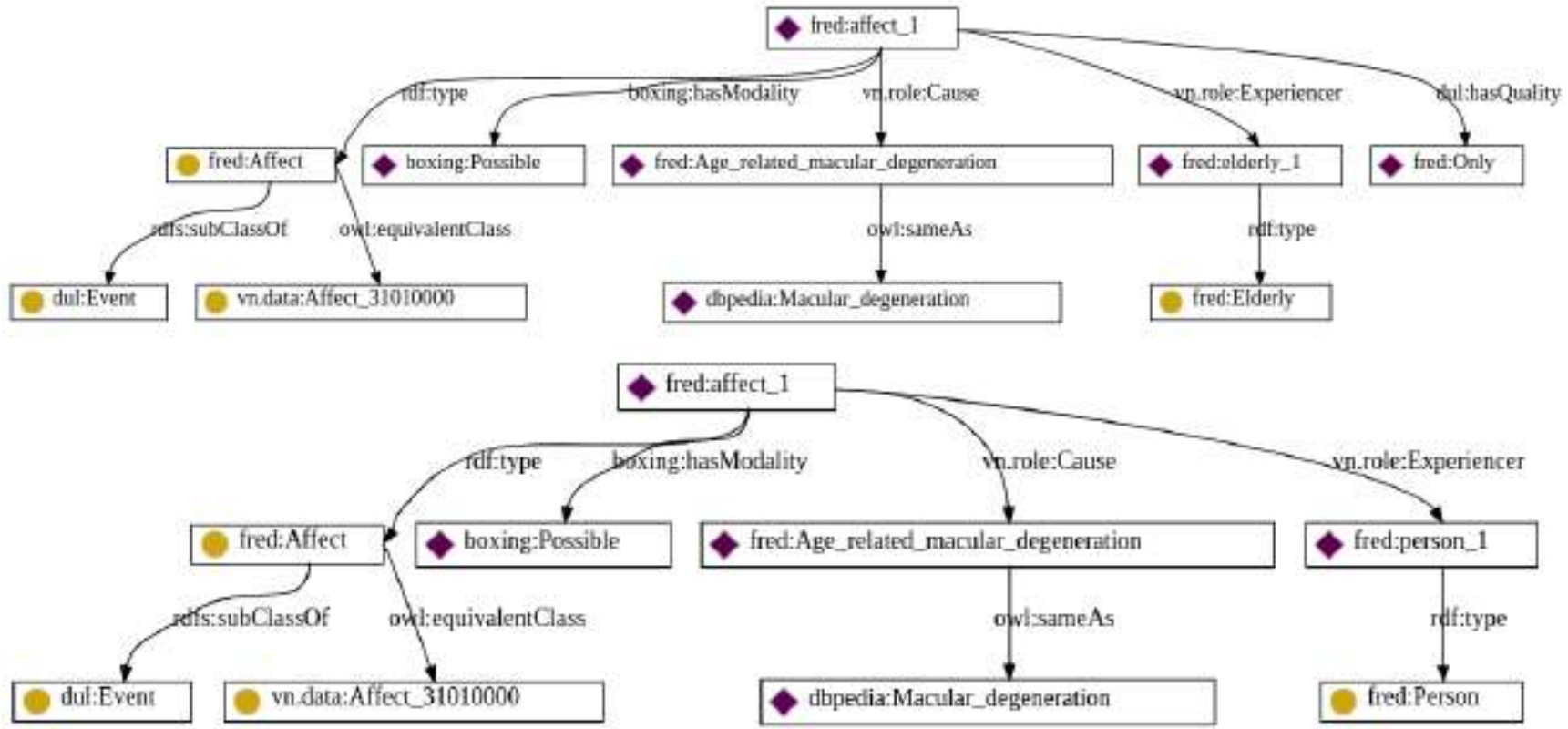


Detecting myths on retinal conditions

Example (Eye and Vision Myths)

- Babies are born with their eyes fully-grown.
- All babies are born with blue eyes.
- Losing vision is an inevitable part of aging.
- Eyes can be transplanted.

Myth: AMD can affect elderly only: $AMD \sqsubseteq \forall affects.Elderly$
Fact: AMD can affect anyone: $AMD \sqsubseteq \forall affects.Person$
Abox: $affectedBy(noey, AMD)$
$hasAge(noey, 50)$
Background knowledge: $Elderly \sqsubseteq Person \sqcap (> hasAge\ 65)$
$affects^{-} \equiv affectedBy$
$AMD \equiv \{amd\}$



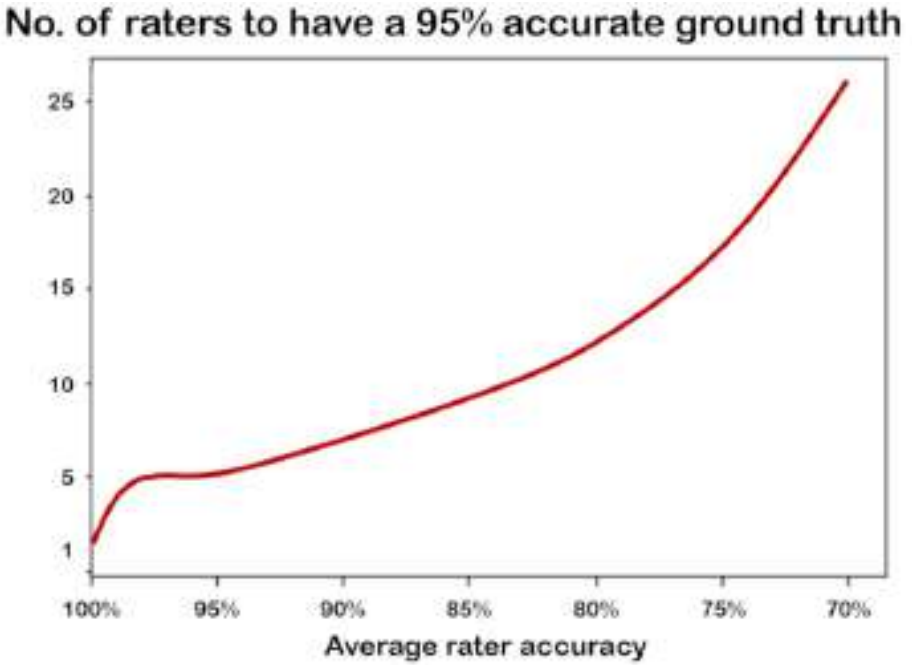
Detecting inconsistency/incoherence by reasoning in Description Logics
Automatic counterspeech generation by verbalising the inconsistency

Ongoing work

Predictive reliability: measuring uncertainty in a single instance
Ground truth data are (wrongly) considered 100% accurate:

Example (The Elephant in the Machine (Cabitza et al.))

- in diagnostic, the average accuracy of medical experts ranges from 80% to 90%
- the average error rate among radiologists is around 30



Applying the metric developing for assessing the case dificulty for residency training to the task of predicting when the model is unreliable

In line with XAI: advocates on the importance to advise users when the model's predictions may be unreliable.
Risk management is mandatory for AI medical applications

Results (based on neuro-symbolic AI)

- Predicting disease evolution from small-time series
- Explaining decision to ophtalmologist and building assurance cases
- A neuro-symbolic case allocation algorithm for residency training
- Formalising AMD diagnosis protocol in Description Logics
- Segmentation of biomarkers from OCT images
- Signaling myths on retina and counterspeech generation (ongoing work)
- Identifying new biomarkers for AMD (ongoing work)

Research Team

- University of Medicine and Pharmacy Iuliu Hatieganu: Simona Delia Nicoara, George Muntean, Ioana Damian, Andrada Dragan, Corina Suciu
- Technical University of Cluj-Napoca: Adrian Groza, Anca Marginean, Radu Slavescu, Raluca Brehar, Pop Adrian

$\forall x\ participates(x, thisSession) \rightarrow thank(I, x)$

ISI articles (since 2021)

1. Groza A, Toderean L, Muntean G. A., Nicoara D. Agents that argue and explain classifications of retinal conditions. Journal of Medical and Biological Engineering. 2021 Oct;41(5):730-41
2. Marginean B. A., Groza A., Muntean G., Nicoara S.D. Predicting Visual Acuity in Patients Treated for AMD. Diagnostics. 2022 Jun 20;12(6):1504
3. Bilc, S.; Groza, A.; Muntean, G.; Nicoara, S.D. Interleaving Automatic Segmentation and Expert Opinion for Retinal Conditions. Diagnostics 2022, 12, 22.
4. Cheres. I., Groza., A "The Profile": unleashing your deepfake self, Multimedia Tools and Applications, Multimedia Tools and Applications, 2023
5. Muntean G. A., Groza A., Marginean A., Steiu M., Muntean V., Nicoara S. D. Artificial intelligence for personalized ophthalmology residency training, J. of Clinical Medicine.
6. Marginean A. N., Muntean D. D., Muntean G. A., Priscu A., Groza A., et al. Reliable Learning with PDE-Based CNNs and DenseNets for Detecting COVID-19, Pneumonia, and Tuberculosis from Chest X-Ray Images. Mathematics. 2021; 9(4)