Microsoft® Research

Faculty Summit 2010

Guarujá, Brasil | May 12 – 14 | In collaboration with FAPESP

Microsoft® Research
Faculty Summit 2010
Guarujá, Brasil | May 12 – 14 | In collaboration with FAPESP

Automatic detection of diabetic retinopathies

Jacques Wainer (wainer@ic.unicamp.br)

Associate Professor – Computer Science - UNICAMP
Visiting Professor – Medical Informatics - UNIFESP

Research team

UNICAMP

Faculty:

- Anderson Rocha
- Siome Goldenstein

Students:

- Tiago Carvalho
- Bruno Brandao

RECOD (reasoning on complex data)

UNIFESP

Faculty:

- Ivan Pissa
- Michel Farah
- Paulo Shor

Specialist:

- Eduardo Dib

Student:

- Fernando Sousa

Health Context

Type 2 Diabetes Mellitus is a growing problem for developed and developing countries.

- 10.7% of people age 20 or above have DM [1]
- a common complication of DM are the Diabetic Retinopathies – 40% of DM patients have them, 8% are vision-threatening[2]
- DR is the main cause of preventable blindness in the US[1] (but not in Brazil or India - cataract[3])

[1] www.diabetes.org data for US 2007

[2] Archives of Ophthalmology 2004; 122:552-563 data for the US 2004.

[3] World Health Organization Visual Impairment and Blindness

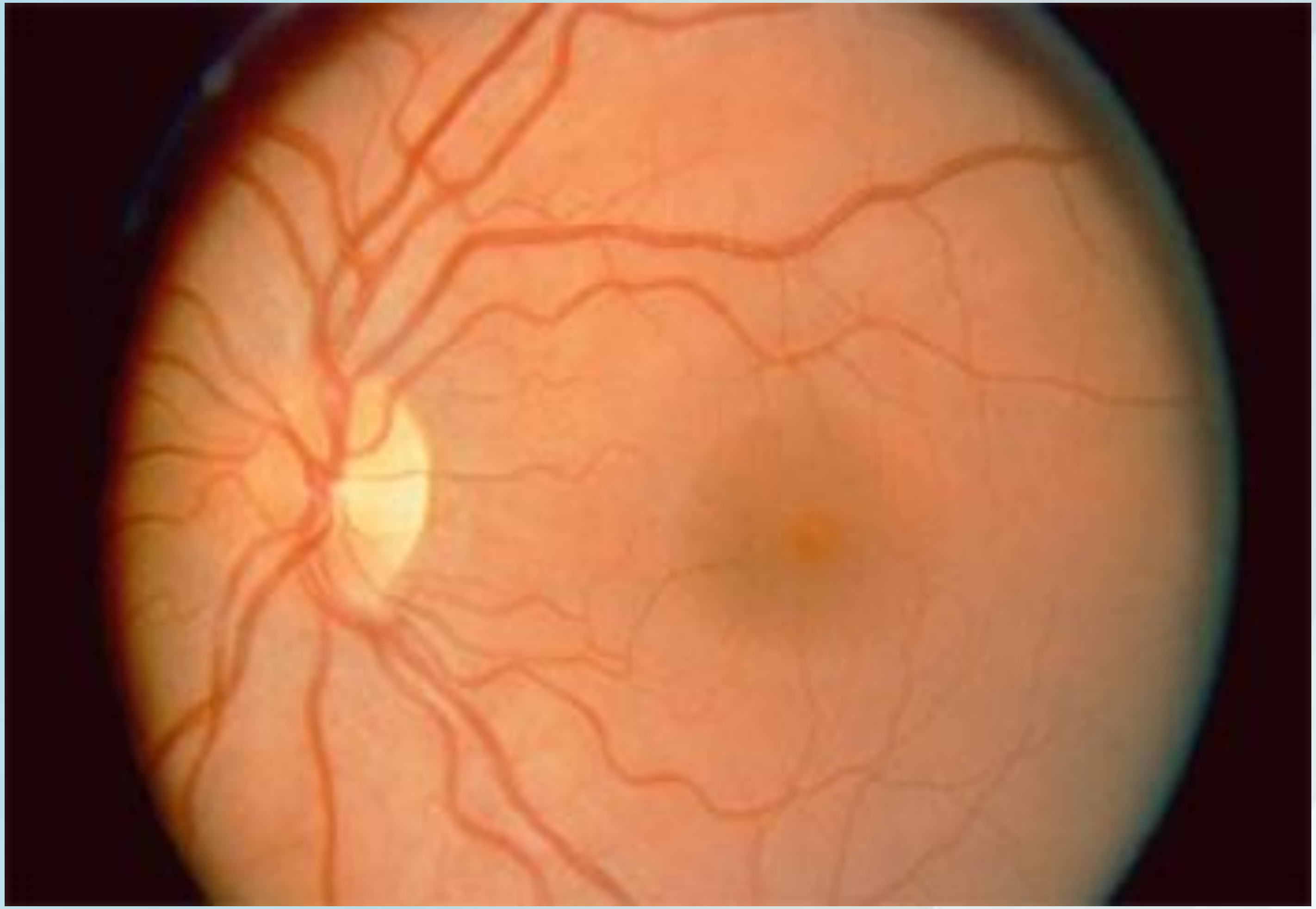
Diabetic Retinopathy

Diabetes destroys the ability of small blood vessels to contain fluids, so there are different forms of leakage.

Problems in organs with very fine vessels – kidney and retina.

Retina – leakage of fat, blood, other liquids may cause different abnormalities

Normal retina



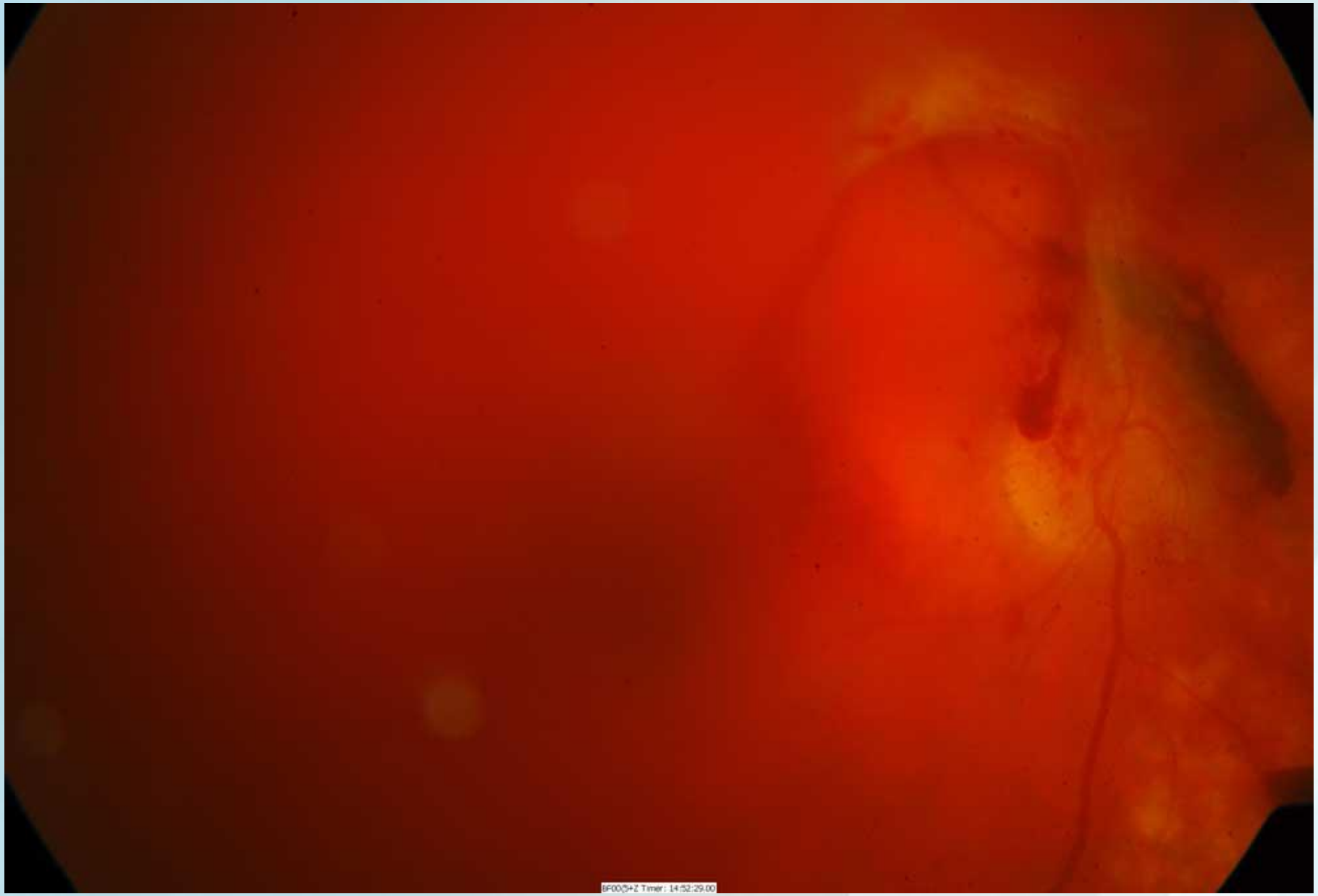
Exudate

Micro aneurysms and small hemorrhages

Proliferative

Proliferative with severe hemorrhages

Hemorrhages into the eye



Public health context (Brazil)

Free public health system is organized in levels of complexity

- Primary care level – Basic Health Unit (UBS) and Family Health Program (PSF)

  general practitioner physicians and nurses (static or home based)

- Secondary level – specialists, out-patient clinics

- High complexity centers – hospitals.

Public health context (Brazil)

- UBS may host a retinograph but VERY unlikely an Opthalmologist
- A health technician may operate a retinograph but only a physician can make any diagnostic regarding retinopathies

Proposal

A automatic tele-ophthalmology project:

- Images taken by a technician at the UBS
- Sent to a system that classifies the image as positive (shows some retinopathy) or negative (normal)
- Positive images are sent to a specialist for diagnostic and possible treatment
- Negative images are NOT sent to anyone

Objectives

1) develop the automatic detection system

2) deploy the system for 6 months in a real tele-opthalmology service linking 2 point-of-care sites
   - an UBS
   - an outpatient clinic for diabetic patients (secondary level)

Questions

- Can an automatic screening program be developed?
- Is such tele-ophthalmology service economically viable?

Requirements of the system

The automatic system **cannot** say that a patient that **has** a problem is a **negative**

- A false negative will not be analyzed by a doctor and the patient may go blind (in the bad case scenario),

Requirements

Hard requirement

- false negatives = 0
- false negative rate = 0
- sensitivity = 100% (health)
- recall = 100% (IR)
- negative predictive value = 100%

or very close to it.

Requirements (II)

And how about a false positive?

- The patient has no problem, but the system flagged the image as positive.
- OK! It only increases the specialist work load – he/she will have to see and analyze images where there is no problem – and may hinder the economic viability of the project

- Soft requirement:
  - false positive rate as low as possible
  - or specificity (health) or precision (IR) as high as possible

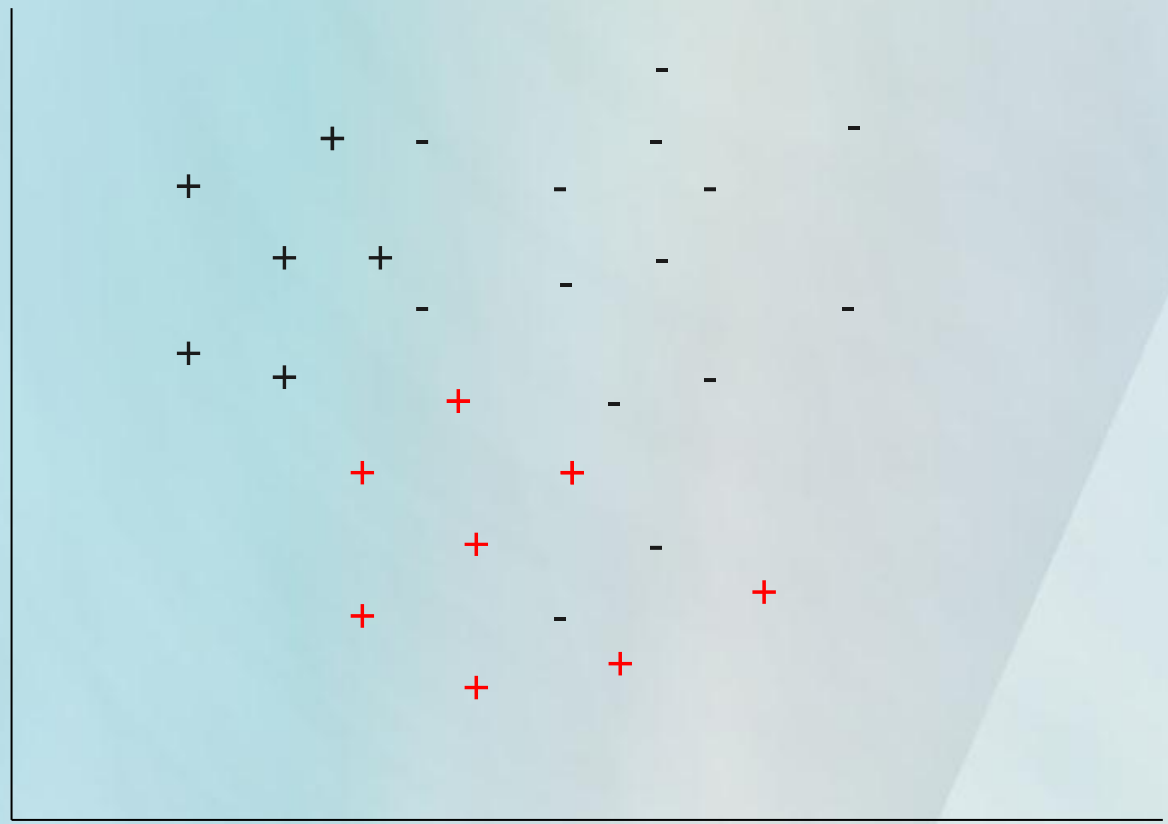
System alternatives

1) Multiple “model specific” specialists for each disease – if any classifier detect a disease, mark the patient as positive

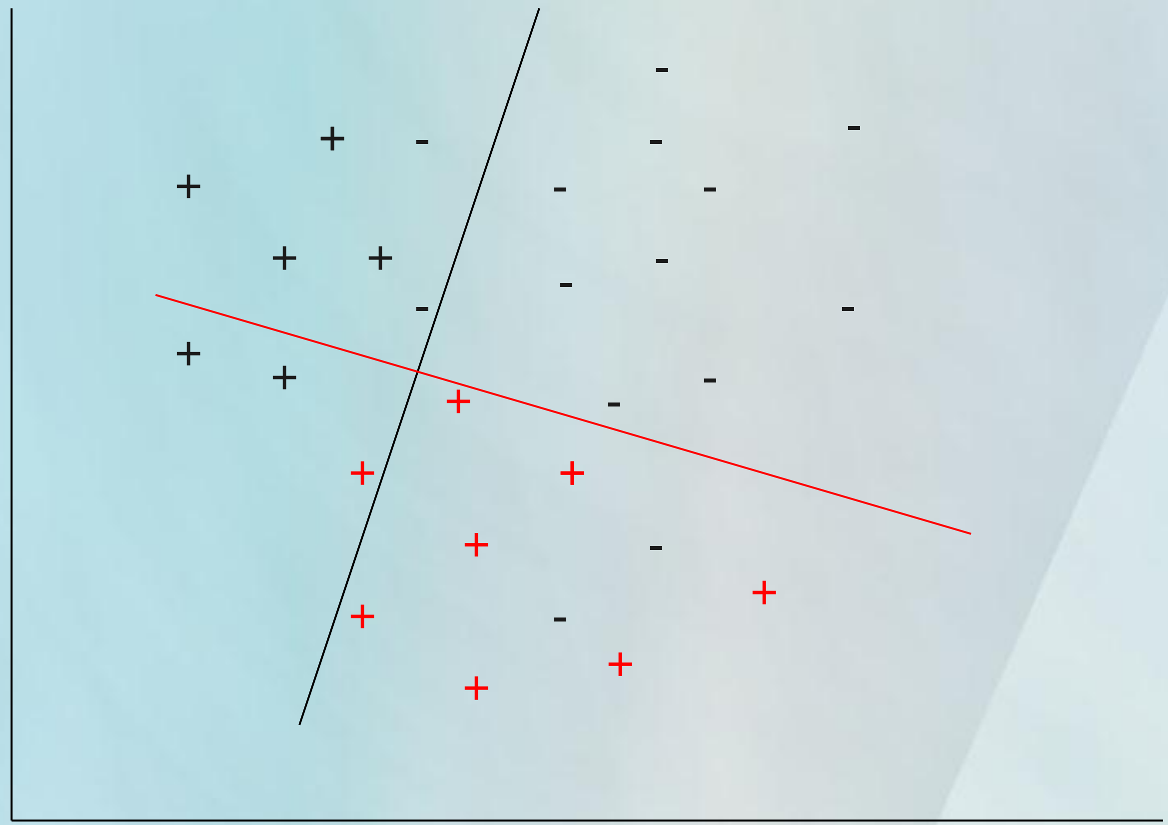
2) Multiple “model-free” specialists of each disease – same as above

3) Learn a classifier for normal/not-normal as a whole – learn the general distinctions between positive and negative examples

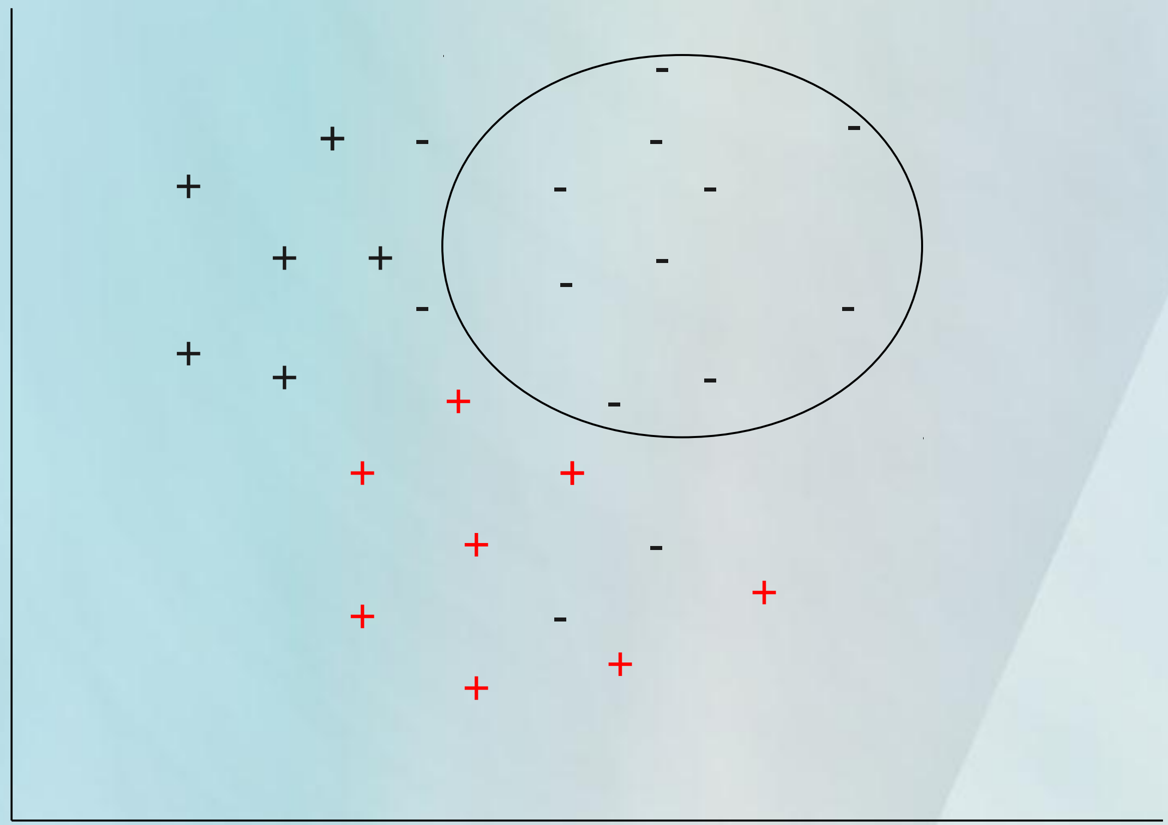
System Alternatives



Multiple specialists (both alternatives)



Normal/not-normal alternative



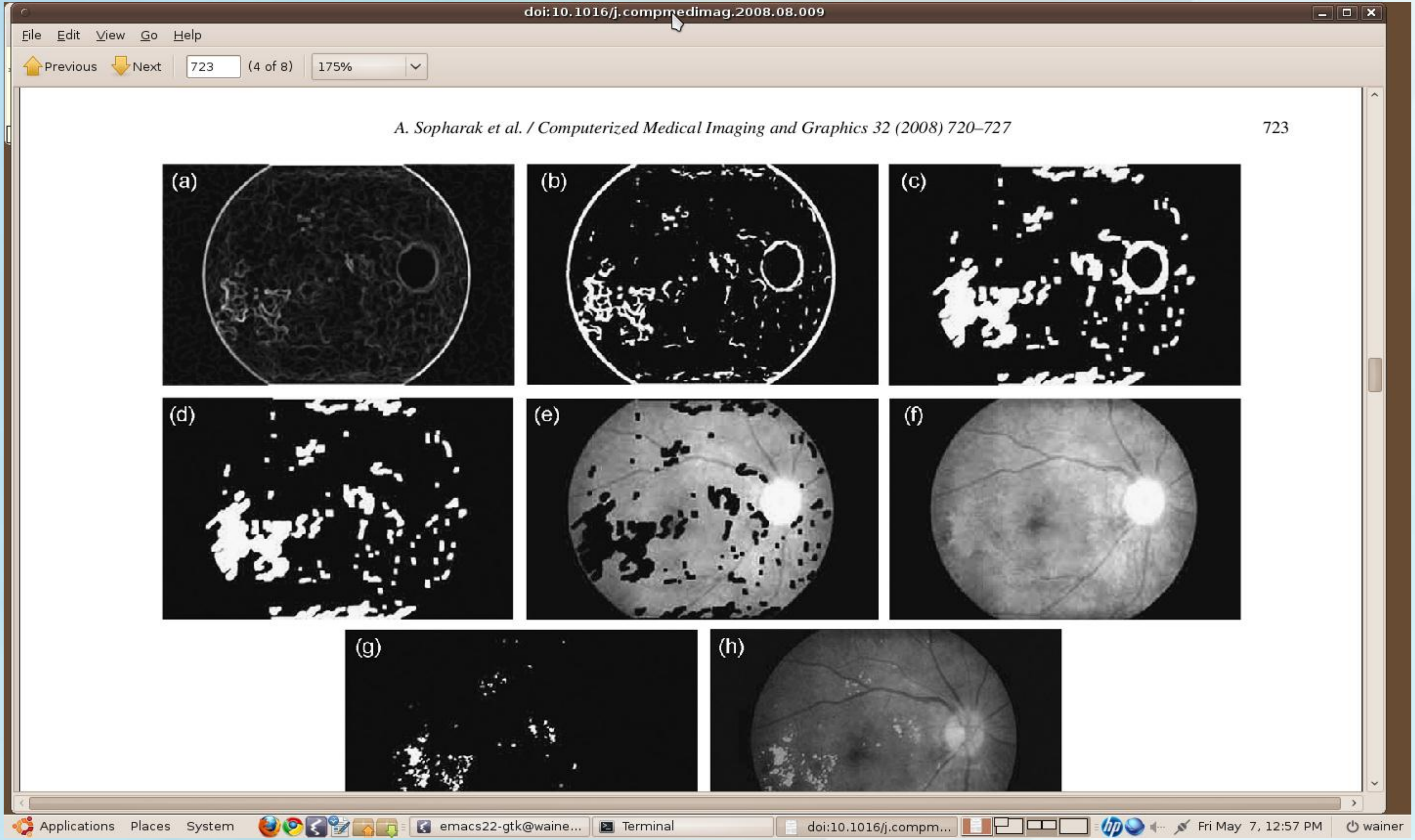
Model specific alternative

- Define with some precision what the abnormality ”looks like”
  - Shape
  - Color
  - Texture
  - Neighborhood

- Devise operations on the image that find candidates

- Select among the candidates

Model specific alternative



Model specific alternative

- Most common approach
- Good accuracy
- Helpful to the specialist – decision support system

but

- Takes long to develop
- Require too much input from the specialist to develop the model of what the abnormality can look like
- Specific to a single anomaly

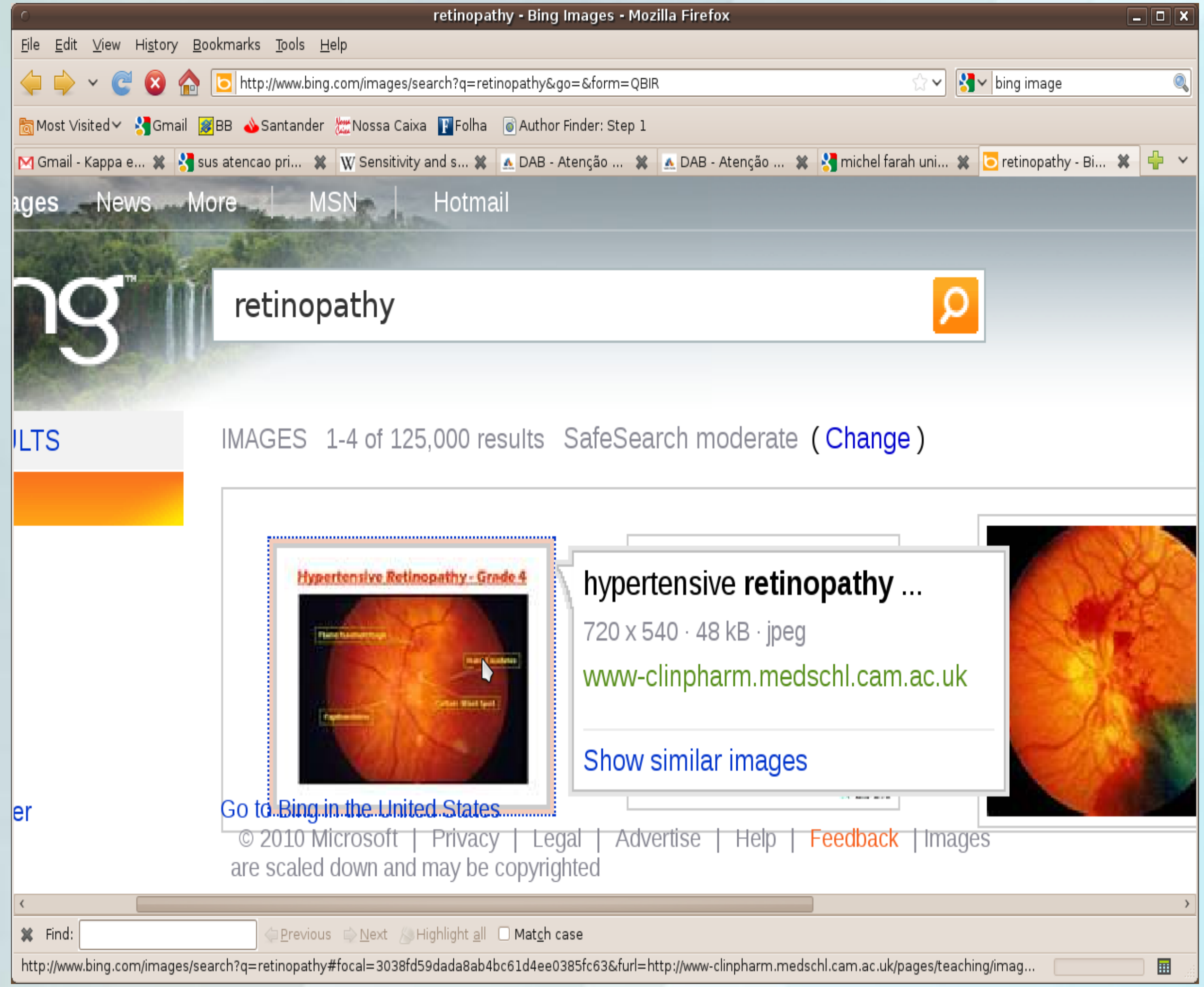
We have 30+ different anomalies and a 2-year project!

Model free alternative

Content based image retrieval (CBIR)

- “show similar images” in image search engines

The system must find similar images. No “model” at programming time.



Model free alternative

- Based on characteristics of keypoints
- We used SURF[1] to select and describe the keypoints.
- Keypoints are points of discontinuities in texture (color, scale, distortion, and orientation invariant)
- Each keypoint has 128 features (besides its location in the picture)
- Each image will have 80 to 2000 keypoints.

[1] H Bay, A Ess, T Tuytelaars, L Gool "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008

Model free alternative

- Descriptions of keypoints (from many images) are grouped in “visual words”
- Each visual word is a “type” of keypoint, and each image will have some of these “types” of keypoints
- The problem of image retrieval becomes similar to text retrieval, using visual words as analogues to text words.
- For our problem – detection – it becomes similar to text classification
- As far as we know this has not been used in retinal image processing.

Model free alternative

Scientific questions:

- Can this technique deal with all abnormalities in diabetic retinopathy?
- Can it achieve the same accuracy of model-specific approaches?

Engineering question:

- Does this technique allow us to develop the different specialist detectors for each of the abnormalities in less than 2 years?

Normal/not-normal alternative

Normal/not-normal alternative

- define some set of visual words that are part of most normal images
- DR image would have small intersection with the set the "normal" visual words.
- **Scientific question**: will this work?

- we had some negative results with Latent Dirichlet allocation[1] for this idea. We are now trying one class SVM[2]

[1] D Blei, A Ng, M Jordan, J Lafferty (January 2003). "Latent Dirichlet allocation". Journal of Machine Learning Research 3: pp. 993–1022.

[2] B. Scholkopf, J.C. Platt, J.Shawe-Taylor, A.J. Smola, and R.C. Williamson. Estimating the support of a high-dimensional distribution. Technical report, Microsoft Research, MSR-TR-99-87, 1999.

Data

Up to now, we have 8,039 images

- They were classified by retina specialists (11)

- Each image was given a very detailed diagnostic (multiple classification) but the anomaly was not marked in the image

- Most frequent abnormalities: exudate (300), deep hemorrhages, increased vascular tortuosity, and druses

Data

- 3306 are non-central (no optic disk or macula)
- 1732 were of poor quality
- 634 vitreous opacity

- 687 normal
- 1694 with some retinopathy

The 2381 good quality images and their classifications will be in public domain at the end of the project (pending approval by the ethics committee at UNIFESP)

State of the art

Technique	Problem	Sensit.	Specif.	Data set	Research Approach
Fleming et al. (2007)	Exudates vs. Drusen vs. Normal	95%	84,6%	13.219 images (300 exudates)	Multi-scale decomposition and Morphological Operators
Hsu et al. (2001)	Exudates vs. Normal	100%	74,2%	543 images (31 exudates, drusen present)	
Lee et al. (2001)	Exudates vs. Normal	96%	93%	422 images (54 with exudates)	
Li and Chutatape (2004)	Exudates vs. Normal	100%	71%	35 images (28 exudates)	
Niemeijer et al. (2007)	Exudates vs. Normal	95%	86%	300 images (42 exudates, 52 drusen, 30 with cotton wool spots)	
Osareh et al. (2003)	Exudates vs. Normal	93%	94,1%	67 images (27 exudates)	
Sinthanayothin et al. (2002a)	Detect Exudate segments	88,5%	99,7%	60790 segments with 10x10 pixels from 30 images (21 with exudates)	Moat operator
Philips et al. (1993)	Detect pixels belonging to Exu-	87%	92,4%	Pixels in 30 regions of 13 images	

State of the art

Lalond et al. (2004)	Detect Exudates, Microneurysms, Anatomical Structures (e.g., optic disk and macula)	100%	87%	46 images	Image registration
Sopharak et al. (2008)	Exudates vs. Normal	80%	99,5%	60 images (40 exudates)	Morphological operators
Kose et al. (2008)	Segmentation of anatomical structures	90% accuracy		60 images	Image segmentation
Sopharak et al. (2009)	Exudates vs. Normal	87,3%	99,3%	60 images (40 exudates)	Fuzzy C-Means clustering + Morphological operators
Abramoff et al. (2008)	Human specialists vs. Automated system to detect any retinal problem	H(85%) A(84%)	H(89%) A(64%)	7689 images for (A), subset of 500 for (H)	Combination of state-of-the-art retinal problem detectors . Optic disc, retinal vessels, hemorrhages, microaneurysms, vascular, abnormalities, exudates, cotton wool spots, drusen detectors
Philip et al. (2007)	Disease vs. No disease	90,5%	67,4%	1067 training and 14406 testing	No details – Proprietary software. Seems to use morphological operators similar to Fleming et al. (2007)

State of the art

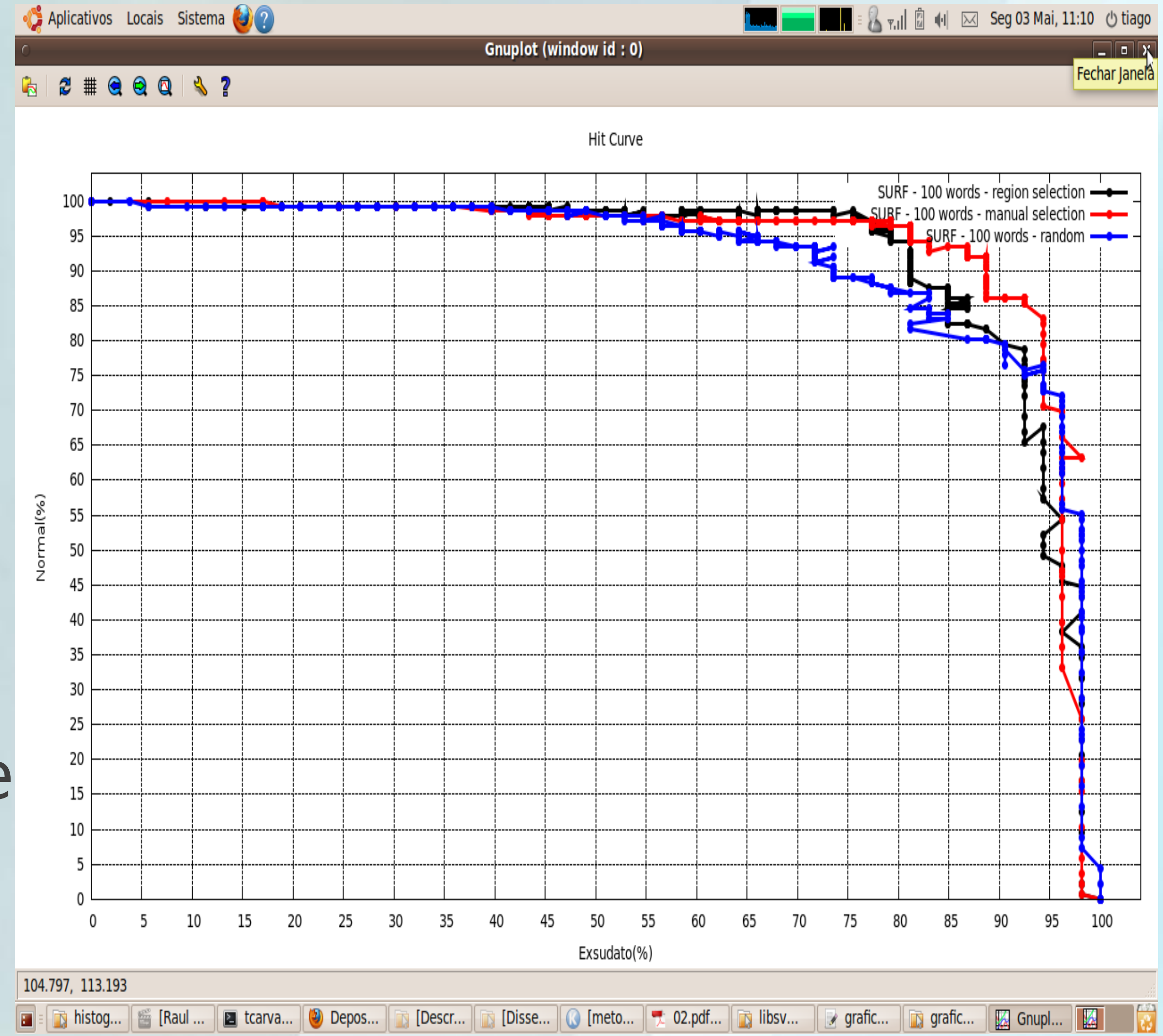
Neubauer et al. (2005)	Disease vs. No disease	93%	100%		Retinal thickness analyzer
Estabridis and Figueiredo (2007)	Disease vs. No disease	90%	accuracy		Identification of the fovea, blood vessel network, optic disk bright and dark lesions
Li et al. (2008)	Disease vs. No disease	81%	Not reported		Bright lesions detection with analysis of retinal vessels patterns
Nayak et al. (2008)	Disease vs. No disease	90%	100%		Analysis of blood vessels, exudates and texture
Acharya et al. (2008)	Disease vs. No disease	83%	89%		Higher order spectra features and support vector machines
Vujosevic et al. (2009)	Grade clinical levels of DR and diabetic macular edema	82%	92%		
Bouhaimed et al. (2008)	Disease vs. No disease	93%	78%	458 images	Retinalyze System (proprietary software)
Garcia et al. (2009)	Hard exudates vs. Normal	88%	84%	117 images (90 with DR)	Neural Networks and support vector machines over patches of images summa-

Results

- no results yet on the normal/not-normal approach
- we started the visual words approach with exudate x normal
  - give us the know how on using this approach
  - most frequent abnormality in our data
  - most frequently used in other papers

Results

- Best result:
  - 95% sensitivity,
  - 85% specificity,
  - 100 visual words
  - no clustering
- Compare with the usual 5000 visual words in CBIR + clustering step (computationally costly)
- Similar to model specific state of the art results



Contributions

- Model free approach to retina image processing – new
- Engineering bet: these ideas can be easily adapted to other abnormalities
- Engineering bet: keypoints can be used in a normal/not-normal approach

Next steps

- normal/not-normal approach
- model free approach to a few other abnormalities – probably there is no need to deal with all of them – there are high correlation among abnormalities
- run the real-life experiment (UBS and out-patient clinic)
  - late according to the plan
  - should start in 2011
  - meanwhile we are collecting new data