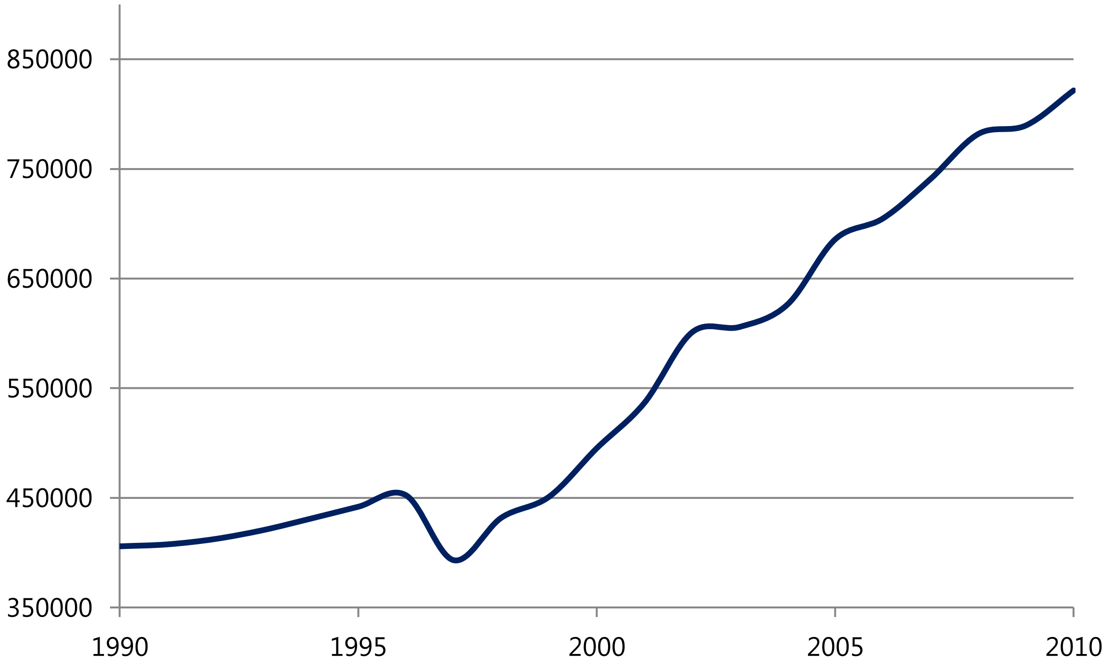
Microsoft Research FacultySummit

Semantic Tagging and its Applications

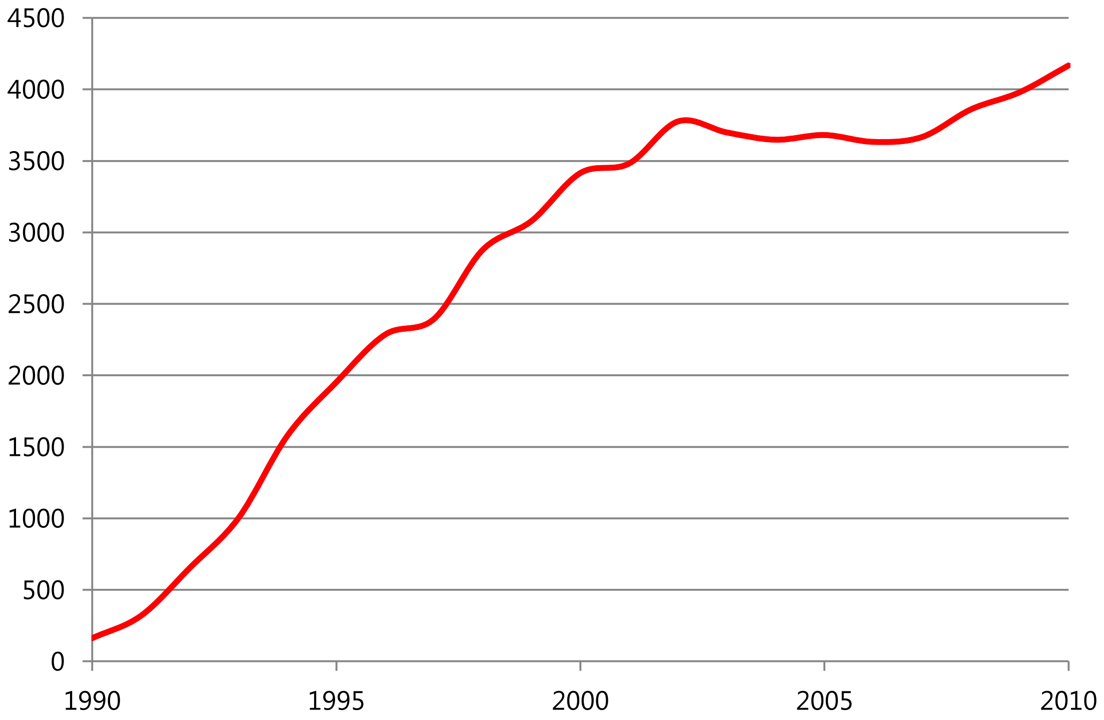
David Haynor
Department of Radiology
University of Washington

FUTURE/WORLD 2011 2031

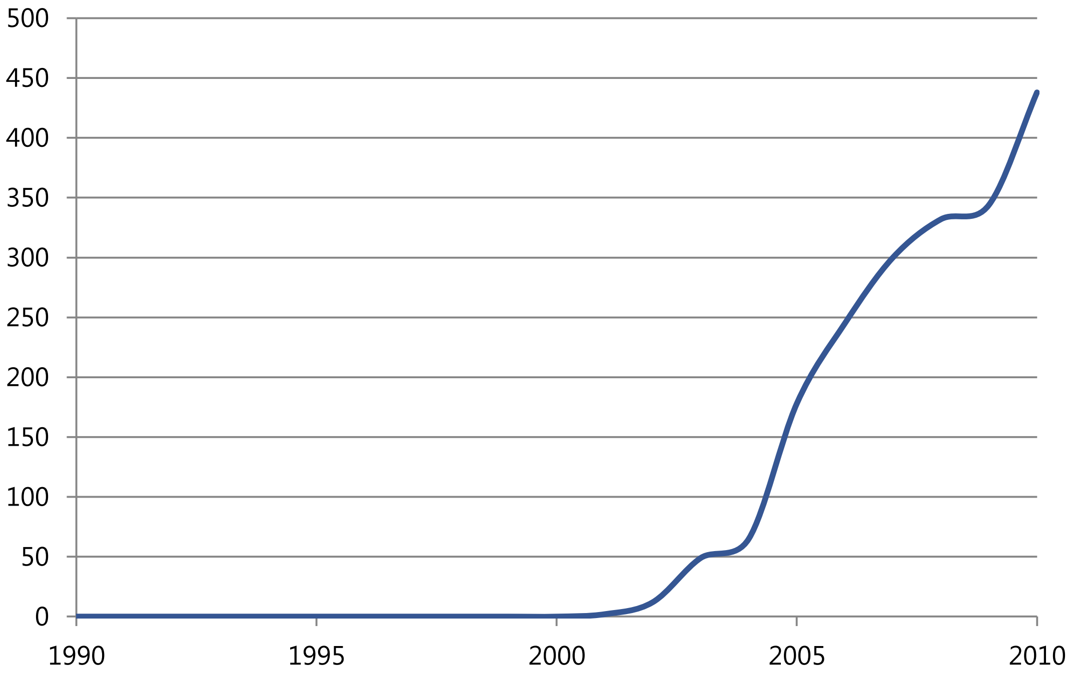
Articles added to PubMed, 1990-2010



Articles on p53, 1990-2010



Articles on erlotinib (Tarceva), 1990-2010



Consequences of information growth in medicine

- New information rate constant-linear with time
- Rate of accrual, even on specialized topics, greatly exceeds practitioner capacity
- Exacerbated by drive to increase productivity
- Productivity, not optimality of care, determines reimbursement
- 5000-15000 cross sectional images per day typical load
- **New information sources/practices must come with a zero or negative time cost to be acceptable to practitioners**

CT scans

- Measures electron density
- One person = 50-100M cubic millimeters = approximate size of a CT dataset
- On basis of density alone can distinguish air, fat, soft tissue, calcification/bone, and foreign bodies – rest must be understood by reference to human anatomy
- HMC (400 beds): applicable to 120 CT scans per day

Human variability in 2D

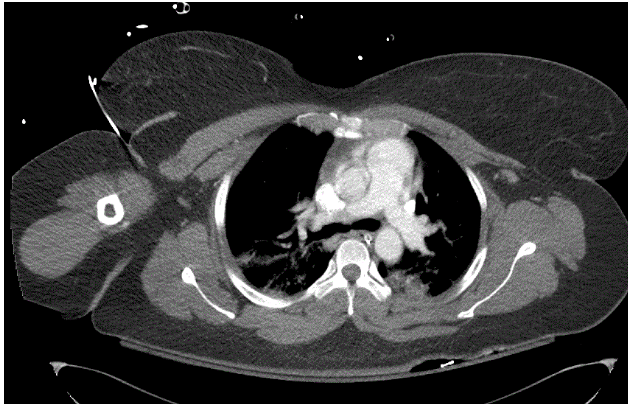
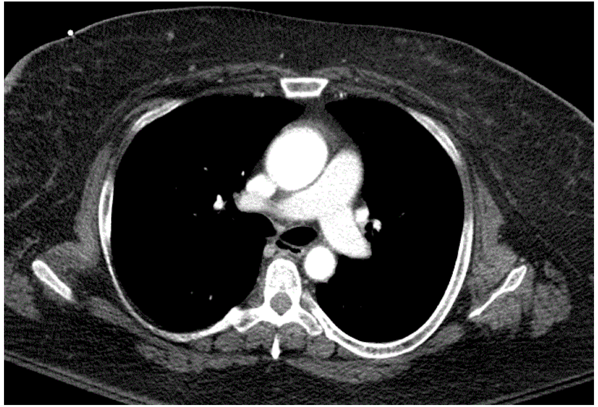
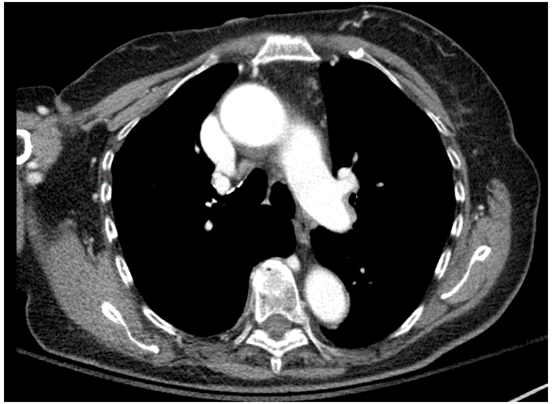
Age
Gender
Color of hair, skin, etc.
Lighting
Foreign bodies

Human variability in 3D CT scans

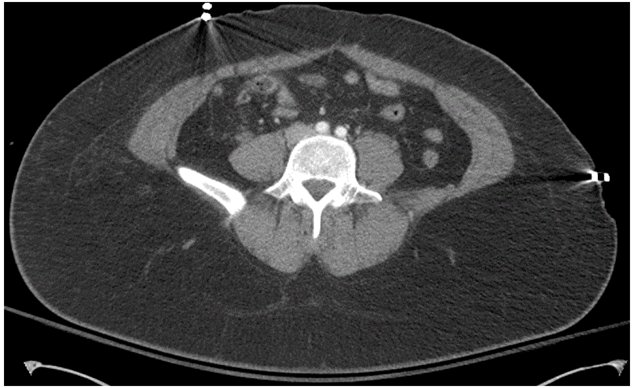
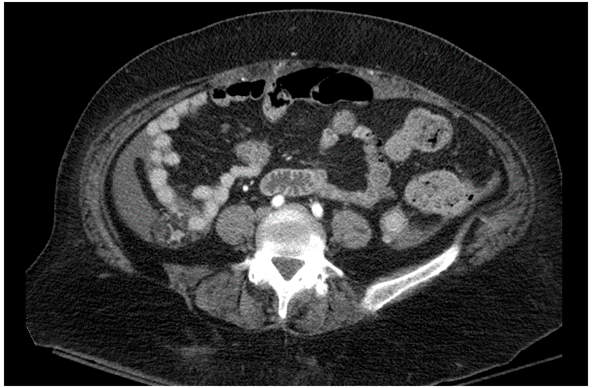
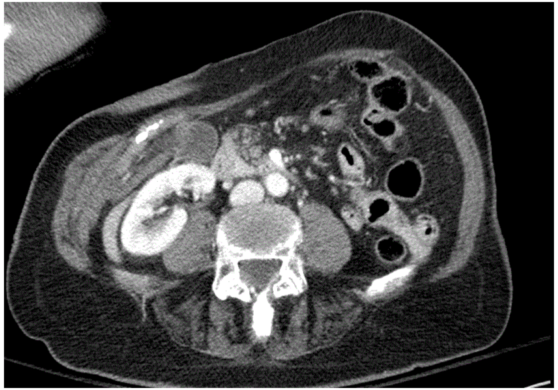
- Age
- Amount of fat vs. muscle
- Variability in size/shape/position/pose of organs
- Gender

- Administration of intravenous/oral contrast

- Effects of disease: organ enlargement/shrinkage/texture change/treatment effects/skeletal deformity

Three women – axial sections at the carina/iliac crests

Human variability: the movie

Semantic tagging is organ labeling

- What voxels are in the (left kidney, liver, femur, T9 vertebra, ...)?
- Historically, the focus was on delineating boundaries, first manually (Live Wire) and then semiautomatically, using texture, intensity, shape, edge models but with initialization left to the user.
- Over the past 2-3 years, researchers have turned their attention to eliminating the manual initialization steps, allowing complete automation of the organ labelling process. **Rough localization** is a recognition or image understanding problem.
- Visual feedback is still required.

Rough localization – state of the art

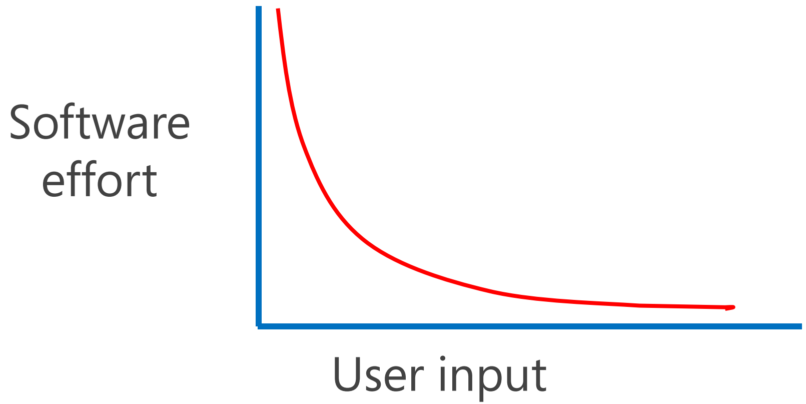
- Liver – Casiraghi et al. IEEE CBMS 2008
- Esophagus – Feulner et al (SCR/MSR) MICCAI 2010 – machine learning
- Heart, great vessels – Ecabert et al. (Philips) Med IA, 2011 model based, machine learning
- Prostate/rectum – Chen et al. (RPI) Med IA, 2011 locate pelvic bones, use model

Rough localization – Criminisi, Kunokoglu

- Goal – localize a bounding box for each organ, using only manually defined training set (so potential to work generally)
- No normalization performed
- Compute many image features at multiple scales at each voxel
- Train multiple regression trees to predict the offset of an organ bounding box from the current voxel using randomized feature sets
- Ensemble prediction using highest-confidence voxels
- Mean error 1-2 cm – runs in seconds
- Generalization to spine – incorporate Markov process

From rough localization to semantic tagging

- Focus of most segmentation studies over past decade
- Combination of characteristic intensities, shape/appearance models, boundary detectors, level sets methods works well -- estimate of bounding box highly informative



Clinical requirements for semantic tagging

- Zero user time
- Display results for quick sanity check
- Feedback/correction

Four applications of semantic tagging

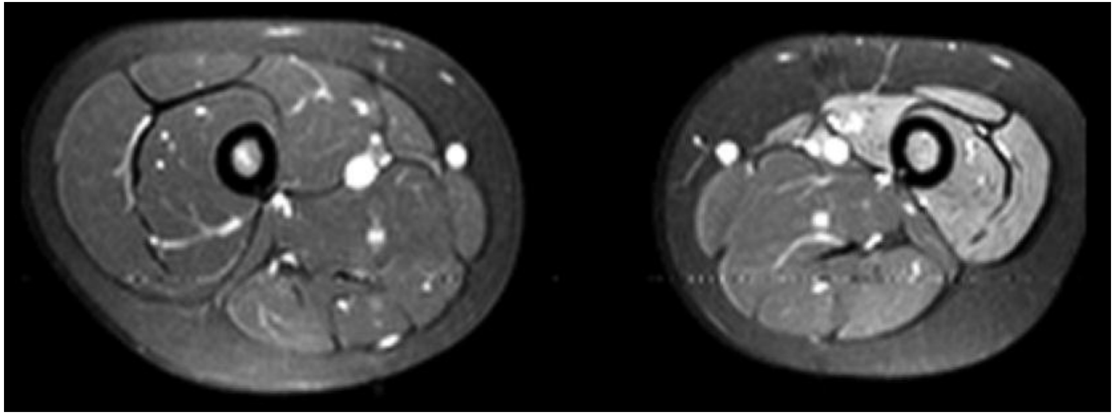
- Semantic tagging turns the image "blob" into a parseable data container with uniquely informative components, much like NLP does with free text

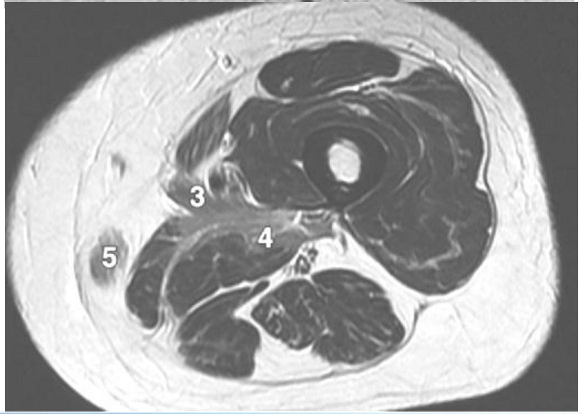
- A deformable atlas of human anatomy
- Detecting changes over time, cross-modality registration
- Osteoporosis and fracture risk
- Quantitative metrics of size, structure, texture: lung, heart, kidney volumes, coronary artery calcification, ...

A deformable atlas of human anatomy

- Semantic tagging, followed by “pretty good” organ segmentation, can give local affine maps between portions of an image set and corresponding portions of an atlas
- These affine maps can be stitched together to form a globally continuous map from image to atlas, so that each point in the image set corresponds to a known point in the atlas.
- Annotated images from the atlas can now be used to automatically annotate the patient’s images.
- Precision of the registration need not be perfect.

What nerves have been injured in these patients?

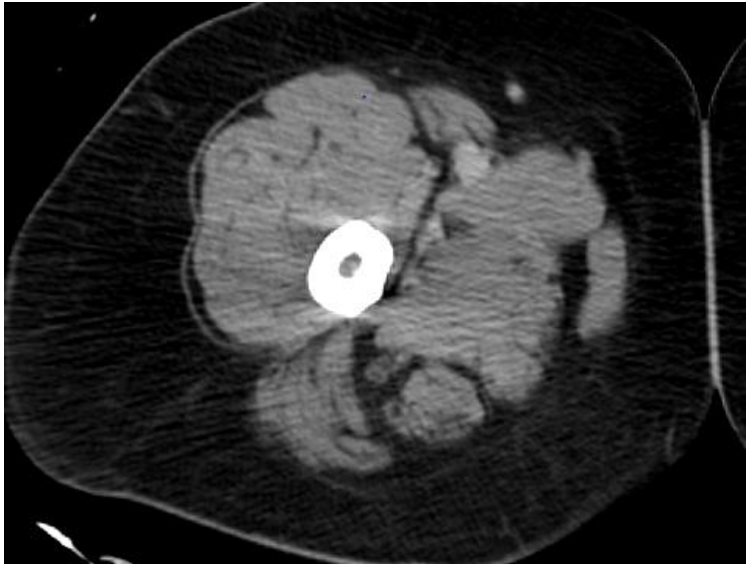




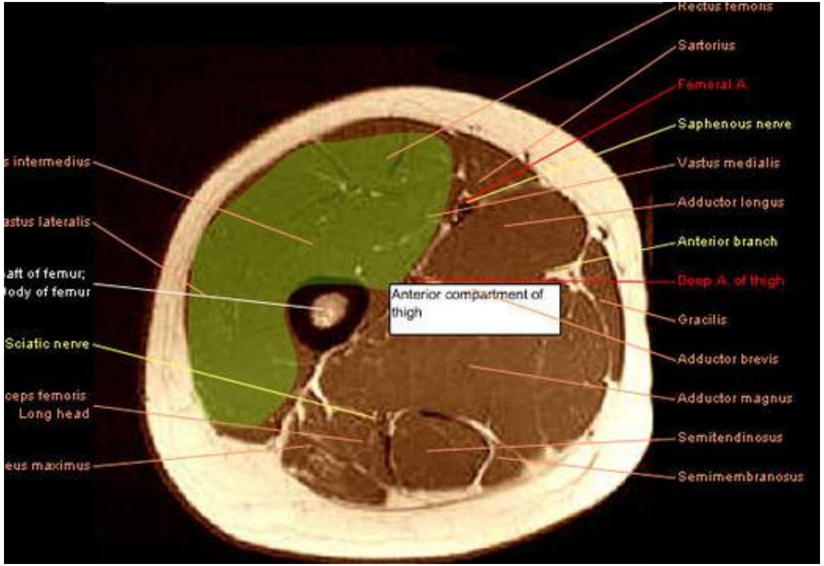
Two patients with leg weakness: 9 yo girl with history of hernia repair; 49 yo woman with history of ovarian cancer.

Questions: (a) where are we; (b) what muscles are affected; (c) which nerve innervates them all?

CT and atlas images



CT through upper right thigh



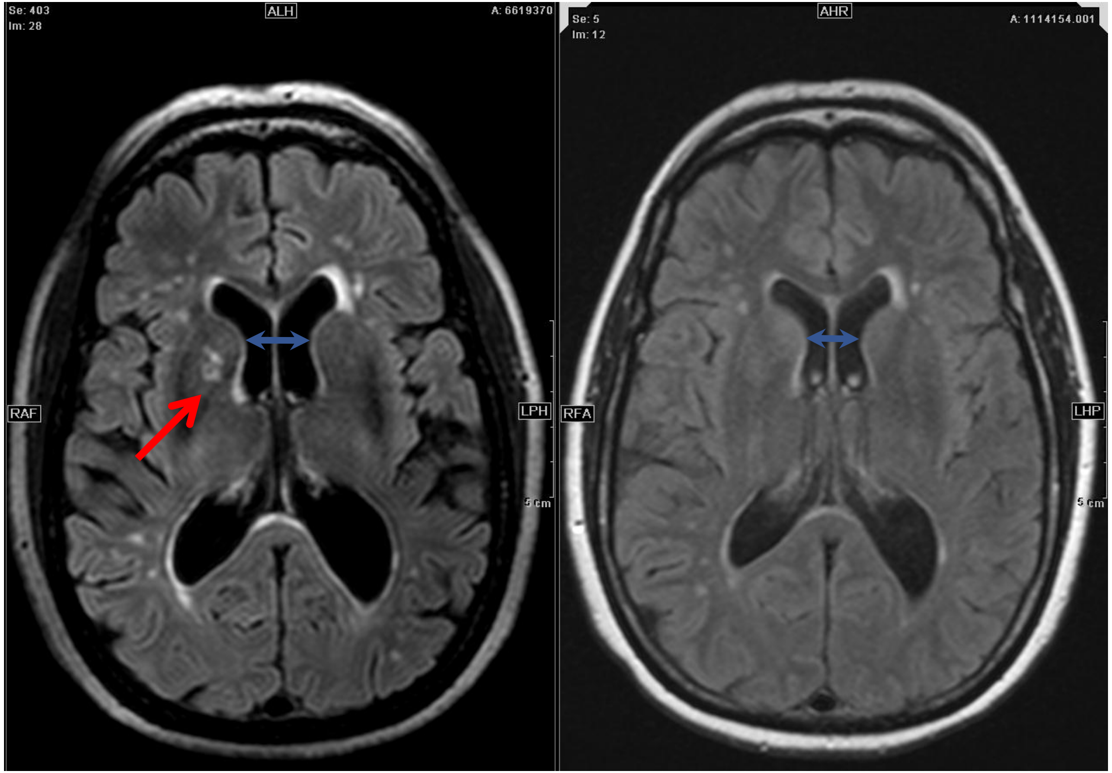
Atlas image (www.imaios.com)

CT and atlas images -- note differences in amount of fat. With semantic tagging, a click on the left image could show the corresponding point on the right image.

Registration and semantic tagging

- Both intramodality (across time) and intermodality registration algorithms within the same patient are highly developed and successful in most cases
- Image comparison is often a much more tedious and error-prone task than diagnosis
- Manual initialization is often required, **breaking the zero cost rule**
- Rough localization could replace the manual initialization step, facilitating comparison of images over time.

57 yo woman with multiple sclerosis



Left: 2011, image 403/28
Right: 2007, image 5/12

There is appearance of new lesions (red arrow) and subtle interval volume loss (blue arrows).

Rough initialization can facilitate image comparison

- Because the initialization is done algorithmically, registration can be performed before the images are viewed.
- Linked cursor or image warping can be used to facilitate visual comparison
- Subtraction of appropriately normalized images can highlight areas of (possible) change

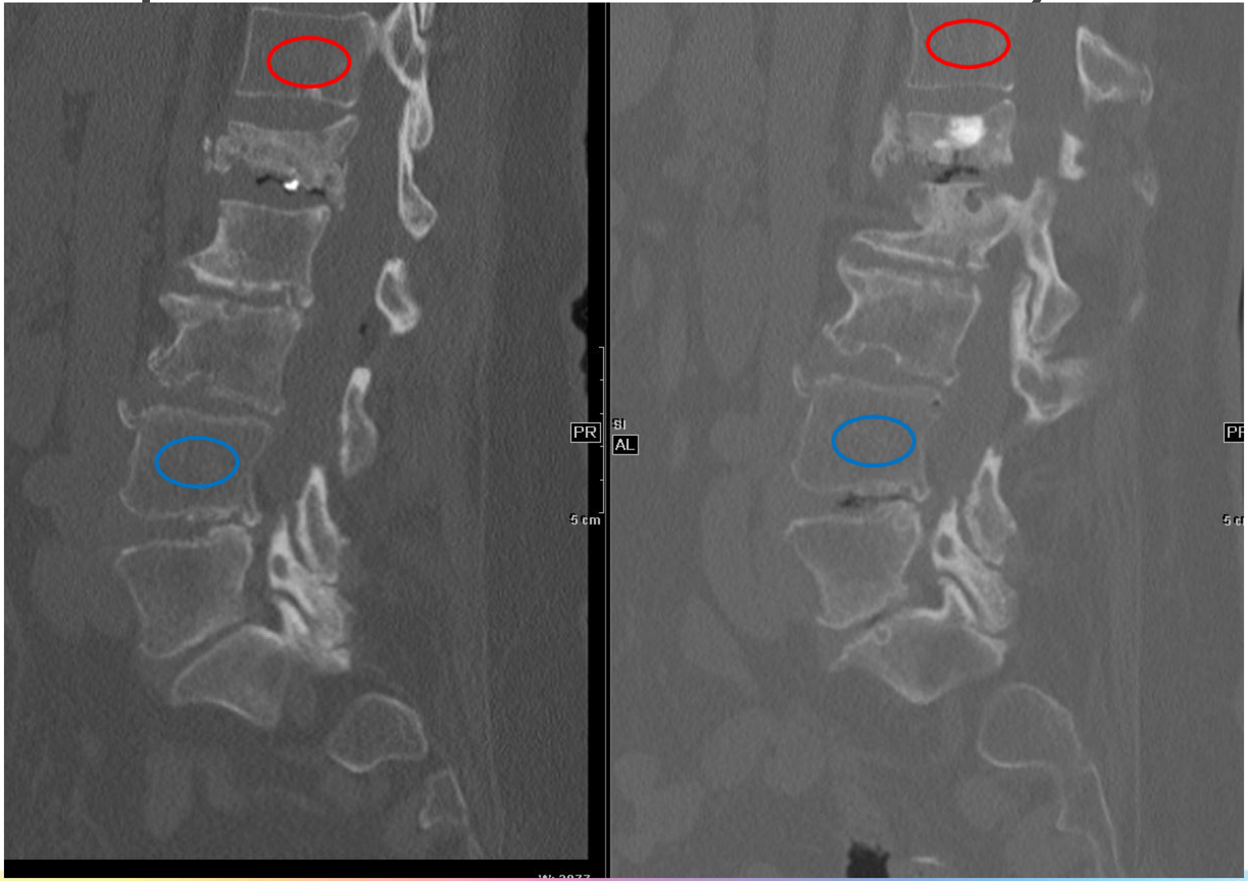
Osteoporosis – decreased bone mass

- 1.5M vertebral fractures per year (US), 180K NHP, $18B
- Immobility, pain, mortality
- Preventable (diet, exercise, quit smoking)
- Treatable (medication, fall prevention, etc.)
- Underdiagnosed
- DXA screening has low compliance rate
- Frequently missed by radiologist on CT

Osteoporotic fracture in a 66 yo woman

Density at time of fx:

T12 = 105
L4 = 81



Density 5 months later:

T12 = 109
L4 = 54

Osteoporosis and semantic tagging

- Identify patients with low bone density for additional evaluation -- density by vertebra, thickness of cortical bone, aggregate indices, …
- Understand results of treatment (mobilization vs. rest, medication, etc.)
- ST makes screening possible at essentially no additional cost to the healthcare system

Quantitative metrics: density/texture

- Liver – fatty liver; cirrhosis
- Bone – osteoporosis
- Lens – cataracts
- Lung – COPD
- Blood vessels – atherosclerotic calcifications
- (MRI applications)

Quantitative metrics: size

- Heart – heart failure, wall thickness
- Lung – asthma/chronic bronchitis
- Liver – cirrhosis, hypertrophy
- Kidney – early renal failure
- Muscle – injury, training effects

Future work

- Improvements in rough localization algorithms (normalization, patient subsetting, increased numbers…)
- Extensions to other body parts (neck, brain, extremities)
- Validation on image sets with significant pathology
- How to represent/store the results of ST?
- New applications and demonstration of clinical utility

Conclusions

- Semantic tagging is the “last mile” of medical image segmentation.
- Significant progress has been made in the last 2-3 years.
- When medical image sets can be parsed, a rich variety of information becomes available that may assist in the management of chronic disease and the early detection of organ damage.

Acknowledgements

- Antonio Criminisi (MSR)
- Ender Konukoglu (MSR)
- Sayan Pathak (HSG)
- Steve White (HSG)

Microsoft® Research
FacultySummit

2031
FUTURE/WORLD
2011