# IMAGE QUALITY

#### AUROBRATA GHOSH RYUTARO TANNO & DANIEL C ALEXANDER

CENTRE FOR MEDICAL IMAGE COMPUTING, UNIVERSITY COLLEGE LONDON

COBCOM WINTER SCHOOL 2017, JUAN-LES-PINS, FRANCE



#### MOTIVATION

#### Human Connectome **Project (HCP)**

1.25mm isotropic, HCP 3T Skyra



#### Research Scanner

- High resolution and SNR
- Long acquisition times
- Expensive



#### MOTIVATION

#### Human Connectome **Project (HCP)**

1.25mm isotropic, HCP 3T Skyra



#### Hospital Scanner

**2.4**mm isotropic, GE





- High resolution and SNR
- Long acquisition times
- Expensive

- Clinical Scanners
- Low resolution and SNR
- Time and cost pressure
- Subsequent analysis affected

#### MOTIVATION

#### Human Connectome **Project (HCP)**

1.25mm isotropic, HCP 3T Skyra



#### Hospital Scanner

2.4mm isotropic, GE





- High resolution and SNR
- Long acquisition times
- Expensive

#### IQT: Machine Learning + Information propagation

#### **Clinical Scanners**

- Low resolution and SNR
- Time and cost pressure
- Subsequent analysis affected



#### ENHANCE DECISIONS IN DEVELOPING NATIONS

#### Developing nations (80% of world's population) rely on older low-field MRI







#### ENHANCE DECISIONS IN DEVELOPING NATIONS









#### ENHANCE DECISIONS IN DEVELOPING NATIONS





#### Enable clinical decisions in developing nations by augmenting low-power MRIs



#### ENHANCE DECISIONS IN DEVELOPING NATIONS





#### DATA HARMONISATION



Scanner-1



Scanner-2



#### Target Scanner



 Enable large scale multi-centre studies and acquisitions

 Normalise data across various scanner models, makes and acquisition parameters

 Enable re-use of old data from phased out scanners

• Facilitate longitudinal studies



#### PARAMETER MAPPING

• Estimate multi-shell model parameters from single-shell data

faster acquisition

 apply on historical single-shell datasets







#### SUPER-RESOLUTION

#### Enhancing diffusion MRI maps: DTI, MAP-MRI, Tractography

#### High-res gold standard

Low-res input

#### [ALEXANDER ET AL 2014, 2017]

#### Interpolation

#### Random Forest IQT





#### SUPER RESOLUTION

#### Random Forest IQT

Bayesian RF IQT: Introducing Uncertainty

Deep Learning IQT with Uncertainty



#### PIPELINE OVERVIEW

• Training:

Begin with HIGH-quality data (HCP dataset) Synthetically downsample to create paired low-quality / high-quality dataset Learn mapping from low-quality —> high-quality • Testing:

Apply mapping to test (low-quality) data to enhance quality



### Down-sampled







### Down-sampled









### Down-sampled

#### Input

HCP data



Output







### PATCH BASED REGRESSION Down-sampled Original HCP data Output Input↓ 6 x (2n+1)<sup>3</sup> 6 x m<sup>3</sup>







































#### Regression: •Random Forest [ALEXANDER 2017]

•Deep Learning [TANNO MICCAI 2017]



### RANDOM FOREST IQT

https://github.com/ucl-mig/iqt





#### RANDOM FOREST IQT [ALEXANDER 2014, 2017]

#### Decision Tree



### RANDOM FOREST IQT [ALEXANDER 2014, 2017] $egin{aligned} \mathcal{D} &= \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \ &= \mathcal{D}_T \cup \mathcal{D}_V \end{aligned}$

#### Decision Tree



# RANDOM FOREST IQT [ALEXANDER 2014, 2017]

#### Decision Tree

(Global Linear)

 $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \quad MX = Y$  $= \mathcal{D}_T \cup \mathcal{D}_V$ 



## RANDOM FOREST IQT [ALEXANDER 2014, 2017]

#### Decision Tree

(Global Linear) M

 $\mathcal{D}_L$  (

 $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \quad MX = Y$  $= \mathcal{D}_T \cup \mathcal{D}_V$ 

4

### Features: $F_1, F_2, ..., F_j \in \mathbb{R}$ Thresholds: $\tau_1, \tau_2, ..., \tau_j \in \mathbb{R}$

 $\mathcal{D}_R$ 



### RANDOM FOREST IQT [ALEXANDER 2014, 2017] Decision Tree

(Global Linear)

#### $\mathcal{D}_L$ Information Gain (training): $I_0 - I_L - I_{R}, \quad I_0 = 2|T| \log \det(S) \\ S = (Y - MX)^T (Y - MX)$

 $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \quad MX = Y$  $= \mathcal{D}_T \cup \mathcal{D}_V$ 

4

### Features: $F_1, F_2, \dots, F_j \in \mathbb{R}$ Thresholds: $\tau_1, \tau_2, \dots, \tau_j \in \mathbb{R}$

 $\mathcal{D}_R$ 



### RANDOM FOREST IQT [ALEXANDER 2014, 2017] Decision Tree

#### $\mathcal{D}_L$ ( Information Gain (training): $I_0 - I_L - I_R, \quad I_0 = 2|T|\log\det(S)$ $S = (Y - MX)^T(Y - MX)$

 $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \quad MX = Y$  $= \mathcal{D}_T \cup \mathcal{D}_V$ 

#### $\mathcal{D}_R$

#### Split test (validation):





### RANDOM FOREST IQT [ALEXANDER 2014, 2017] Decision Tree

(Global Linear)

#### $\mathcal{D}_L$ Information Gain (training): $\begin{bmatrix} I_0 - I_L - I_{R_{I_0}} & I_0 = 2|T| \log \det(S) \\ S = (Y - MX)^T (Y - MX) \end{bmatrix}$

 $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \quad MX = Y$  $= \mathcal{D}_T \cup \mathcal{D}_V$ 

**4**.....

Features:  $F_1, F_2, \ldots, F_j$ Thresholds:  $\tau_1, \tau_2, ..., \tau_j \in \mathbb{R}$ 

#### $\mathcal{D}_R$

#### Split test (validation):

$$\mathcal{E}_{LR} < \mathcal{E}_{P}, \begin{cases} \mathcal{E}_{P} &= \sum_{1}^{|V|} ||\mathbf{y}_{i} - M\mathbf{x}_{i}|| \\ \mathcal{E}_{LR} &= \sum_{1}^{|V|} ||\mathbf{y}_{i} - \mathcal{C}(M_{L}, M)| \end{cases}$$



### RANDOM FOREST IQT [ALEXANDER 2014, 2017] Decision Tree $\mathcal{D}_L$ Information Gain (training): $I_0 - I_L - I_R$ , $I_0 = 2|T|\log \det(S)$ $S = (Y - MX)^T(Y - MX)$

Random Forest: 8+ decision trees (Patch library from 8 subjects)

 $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \quad MX = Y$  $= \mathcal{D}_T \cup \mathcal{D}_V$ 

#### $\mathcal{D}_R$

#### Split test (validation):





### RANDOM FOREST IQT [ALEXANDER 2014, 2017] Decision Tree (Global Linear) $\mathcal{D}_L$ Information Gain (training): $I_0 - I_L - I_{R}, \quad I_0 = 2|T| \log \det(S) \\ S = (Y - MX)^T (Y - MX)$

Features up to 23 (DTI): Random Forest: 8+ decision trees Eigenvalues of diffusion tensor (Patch library from 8 subjects) linearity, planarity, sphericity Means of the features over patch

 $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{|\mathcal{D}|} \quad MX = Y$  $= \mathcal{D}_T \cup \mathcal{D}_V$ 

#### Split test (validation): $\mathcal{D}_R$ $\mathcal{E}_{LR} < \mathcal{E}_{P}, \begin{cases} \mathcal{E}_{P} &= \sum_{1}^{|V|} ||\mathbf{y}_{i} - M\mathbf{x}_{i}|| \\ \mathcal{E}_{LR} &= \sum_{1}^{|V|} ||\mathbf{y}_{i} - \mathcal{C}(M_{L}, M_{R})\mathbf{x}_{i}|| \end{cases}$






















### TREE VISUALISED





### TREE VISUALISED









### TRACTOGRAPHY [ALEXANDER 2017]



- Extension from DTI to MAP-MRI
- Tracing 4 pathways: cortical hand area to:
  - thalamus
  - brainstem
  - spinal cord
  - putamen

 Tractography separates tracts
 in 1.25mm but not in low-res (2.5mm) or linear/cubic interpolation
 but again in IQT super-res



### TRACTOGRAPHY [ALEXANDER 2017]







## RESULTS — SUPER-RESOLUTION



### [ALEXANDER 2017]







BAYESIAN RF IQT: INTRODUCING UNCERTAINTY



## (LOCALLY) BAYESIAN RE IQT [TANNO 2016]

Uncertainty estimation from a (locally) Bayesian inference

• Bayesian linear model at each node:

 $\mathbf{y} = M\mathbf{x} + \eta \qquad P(M_{|}|\alpha) = \mathcal{N}(M_{|}, \alpha^{-1}I)$  $P(\eta|\beta) = \mathcal{N}(\eta, \beta^{-1}I)$ 



## (LOCALLY) BAYESIAN RF IQT [TANNO 2016]

Uncertainty estimation from a (locally) Bayesian inference

Bayesian linear model at each node:

Predictive variance for uncertainty quantification:

 $\sigma_{\text{Pred}}^2(\mathbf{x}^*) = \mathbf{x}^{*'I'} \mathbf{A}(\mathcal{D}) \mathbf{x}^* + \beta^{-1}$ 

 $\mathbf{y} = M\mathbf{x} + \eta$   $P(M_{|}\alpha) = \mathcal{N}(M_{|}, \alpha^{-1}I)$  $P(\eta|\beta) = \mathcal{N}(\eta, \beta^{-1}I)$ 





## (LOCALLY) BAYESIAN RF IQT [TANNO 2016]

Uncertainty estimation from a (locally) Bayesian inference

Bayesian linear model at each node:

Predictive variance for uncertainty quantification:

 $\sigma_{\text{Pred}}^2(\mathbf{x}^*) = \mathbf{x}^{*'I'} \mathbf{A}(\mathcal{D}) \mathbf{x}^* + \beta^{-1}$ 

distance from training data

### $\mathbf{y} = M\mathbf{x} + \eta$ $P(M_{|}\alpha) = \mathcal{N}(M_{|}, \alpha^{-1}I)$ $P(\eta|\beta) = \mathcal{N}(\eta, \beta^{-1}I)$

Noise variability in training data 17













 $M = YX^T \left( XX^T + \frac{\alpha}{\beta}I \right)^{-1}$ 







 $M = YX^T \left( XX^T + \frac{\alpha}{\beta}I \right)^{-1}$ 

### Uncertainty



Uncertainty correlates with accuracy



## UNCERTAINTY

## Multiple Sclerosis

T2



**H** 









### Tumour (edema)









DEEP LEARNING IQT WITH UNCERTAINTY



## BASELINE NETWORK [TANNO MICCAI 2017]

### ESPCN = Efficient Subpixel Convolutional Network [SHI CVPR 2016]

### low-res



### high-res



## BASELINE NETWORK [TANNO MICCAI 2017]

### ESPCN = Efficient Subpixel Convolutional Network [SHI CVPR 2016]

### low-res



### high-res



## BASELINE NETWORK [TANNO MICCAI 2017]

### ESPCN = Efficient Subpixel Convolutional Network [SHI CVPR 2016]

### low-res



layers:  $(3,3,3,50) \longrightarrow (1,1,1,100) \longrightarrow (3,3,3,r.r.r.c)$ (width, height, depth, channels)

### high-res

### 3D extension of ESPCN



### Intrinsic Uncertainty

low-res, x

high-res, y



### Parameter Uncertainty





### Intrinsic Uncertainty



### Heteroscedastic noise model (no parameter sharing)



### Parameter Uncertainty

$$\frac{1}{N} \sum_{i}^{N} ||\mathbf{y}_{i} - \mu_{\theta_{1}}(\mathbf{x}_{i})||_{\Sigma_{\theta}}^{2} + \frac{1}{N} \sum_{i}^{N} \log \det \Sigma_{\theta_{2}}(\mathbf{x}_{i})|_{\Sigma_{\theta}}^{2}$$

![](_page_57_Picture_11.jpeg)

### Intrinsic Uncertainty

low-res, x

high-res, y

![](_page_58_Figure_4.jpeg)

### **Parameter Uncertainty**

![](_page_58_Figure_9.jpeg)

 Approximate (intractable) posterior with a factored Gaussian distribution  $p(\theta | \mathcal{D}) \sim q_{\phi}(\theta)$ [KINGMA 2015] using variational dropout where dropout rates are learned

![](_page_58_Picture_12.jpeg)

### Intrinsic Uncertainty

low-res, x

high-res, y

![](_page_59_Figure_4.jpeg)

### Heteroscedastic + Variational dropout

![](_page_59_Figure_6.jpeg)

### **Parameter Uncertainty**

![](_page_59_Figure_9.jpeg)

• Predictive distribution:

$$\int \mathcal{N}(\mathbf{y}; \mu_{\theta_1}, \Sigma_{\theta_2}(\mathbf{x})) \cdot q_{\phi}(\mathbf{x})$$

Mean & Uncertainty (variance) from dropout sampling

![](_page_59_Picture_14.jpeg)

## Interior

![](_page_60_Picture_2.jpeg)

## Exterior

![](_page_60_Picture_4.jpeg)

### Interior

	Unseen HCP cohort: <b>similar</b> demographics, protocol		Unseen Lifespan cohort: <b>different</b> demographics, protocol	
Models	HCP (interior)	HCP (exterior)	Life (interior)	Life (exterio
CSpline	$10.069 \pm n/a$	$31.738\pm$ n/a	$32.483 \pm n/a$	$49.066 \pm n/a$
IQT-RF	$6.974 \pm 0.024$	$23.139 \pm 0.351$	$10.038\pm0.019$	$25.166\pm0.32$
BIQT-RF	$6.972 \pm 0.069$	$23.110\pm0.362$	$9.926 \pm 0.055$	$25.208\pm0.29$
3D-ESPCN(baseline)	$6.378 \pm 0.015$	$13.909\pm0.071$	$8.998 \pm 0.021$	$16.779 \pm 0.10$

## Exterior

![](_page_61_Picture_5.jpeg)

### Interior

	Unseen HCP cohort: <b>similar</b>		Unseen Lifespan cohort: <b>different</b>	
	demograph	ics, protocol	demographics, protocol	
Models	HCP (interior)	HCP (exterior)	Life (interior)	Life (exterio
CSpline	$10.069 \pm n/a$	$31.738 \pm n/a$	$32.483 \pm n/a$	$49.066 \pm n/a$
IQT-RF	$6.974 \pm 0.024$	$23.139 \pm 0.351$	$10.038\pm0.019$	$25.166\pm0.32$
BIQT-RF	$6.972 \pm 0.069$	$23.110\pm0.362$	$9.926 \pm 0.055$	$25.208\pm0.29$
3D-ESPCN(baseline)	$6.378 \pm 0.015$	$13.909\pm0.071$	$8.998 \pm 0.021$	$16.779 \pm 0.10$
Improvement	9%	43%	10%	33%
		Faster!		

![](_page_62_Picture_5.jpeg)

## Exterior

![](_page_62_Picture_7.jpeg)

### Interior

![](_page_63_Picture_2.jpeg)

Exterior

	Unseen HCP cohort: <b>similar</b>		Unseen Lifespan cohort: different	
	demograph	ics, protocol	demographics, protocol	
Models	HCP (interior)	HCP (exterior)	Life (interior)	Life (exterio
CSpline	$10.069 \pm n/a$	$31.738\pm$ n/a	$32.483 \pm n/a$	$49.066 \pm n/a$
IQT-RF	$6.974 \pm 0.024$	$23.139 \pm 0.351$	$10.038\pm0.019$	$25.166\pm0.32$
BIQT-RF	$6.972 \pm 0.069$	$23.110\pm0.362$	$9.926 \pm 0.055$	$25.208 \pm 0.29$
3D-ESPCN(baseline)	$6.378 \pm 0.015$	$13.909\pm0.071$	$8.998 \pm 0.021$	$16.779 \pm 0.10$
Hetero-CNN	$6.294 \pm 0.029$	$15.569\pm0.273$	$8.985 \pm 0.051$	$17.716\pm0.27$
Var.(I)-CNN	$6.354 \pm 0.015$	$13.824 \pm 0.031$	$8.973 \pm 0.024$	$16.633 \pm 0.08$
Var.(II)-CNN	$6.356 \pm 0.008$	$13.846\pm0.017$	$8.982 \pm 0.024$	$16.738 \pm 0.073$
Hetero+Var.(I)	$6.291 \pm 0.012$	$13.906\pm0.048$	$8.944 \pm 0.044$	$16.761\pm0.04$
Hetero+Var.(II)	$\textbf{6.287} \pm \textbf{0.029}$	$13.927\pm0.093$	$8.955 \pm 0.029$	$16.844\pm0.10$

Top models quantify uncertainty

Best

Second

![](_page_63_Picture_5.jpeg)

## UNCERTAINTY PROPAGATION

![](_page_64_Figure_1.jpeg)

### Model: Hetero + Var (I)

200 samples of predicted DTI

![](_page_64_Picture_5.jpeg)

### UNCERTAINTY IN PATHOLOGY

### Clinical image

![](_page_65_Figure_2.jpeg)

![](_page_65_Picture_3.jpeg)

### Uncertainty correlates with pathology

After Super-res

### Uncertainty

high

![](_page_65_Picture_10.jpeg)

# CONCLUSION

![](_page_66_Picture_1.jpeg)

![](_page_67_Picture_1.jpeg)

![](_page_67_Picture_2.jpeg)

• RF IQT:

![](_page_68_Picture_2.jpeg)

![](_page_68_Picture_3.jpeg)

### • RF IQT:

Patch based regression

![](_page_69_Picture_3.jpeg)

![](_page_69_Picture_4.jpeg)

### • RF IQT:

Patch based regression

Can meaningfully segregate brain patches

![](_page_70_Picture_4.jpeg)

![](_page_70_Picture_5.jpeg)

### • RF IQT:

Patch based regression
Can meaningfully segregate brain patches
Able to successfully super-resolve and is generalisable (monkey brain results)

![](_page_71_Picture_4.jpeg)
#### • RF IQT:

Patch based regression
Can meaningfully segregate brain patches
Able to successfully super-resolve and is generalisable (monkey brain results)
Improves tractography



#### • RF IQT:

Patch based regression
Can meaningfully segregate brain patches
Able to successfully super-resolve and is generalisable (monkey brain results)
Improves tractography

• Locally Bayesian RF IQT:



#### • RF IQT:

Patch based regression Can meaningfully segregate brain patches Able to successfully super-resolve and is generalisable (monkey brain results) Improves tractography

#### Locally Bayesian RF IQT:

Better reconstruction accuracy than regular RF IQT

30



#### • RF IQT:

Patch based regression
Can meaningfully segregate brain patches
Able to successfully super-resolve and is generalisable (monkey brain results)
Improves tractography

#### • Locally Bayesian RF IQT:

Better reconstruction accuracy than regular RF IQT

Uncertainty correlates well with reconstruction error

ular RF IQT ruction error



#### • RF IQT:

Patch based regression
Can meaningfully segregate brain patches
Able to successfully super-resolve and is generalisable (monkey brain results)
Improves tractography

#### Locally Bayesian RF IQT:

Better reconstruction accuracy than regular RF IQT

- Uncertainty correlates well with reconstruction error
- Uncertainty flags pathology (MS, tumour) consistently

ular RF IQT ruction error ır) consistently







#### • Deep Learning IQT:





# Deep Learning IQT: Generalises 3D ESPCN (computationally efficient SR network)



• Deep Learning IQT: Generalises 3D ESPCN (computationally efficient SR network) Variational dropout using a dual-network architecture

# Intrinsic and Parameter Uncertainty modelled using Heteroscedastic noise model and



• Deep Learning IQT: Generalises 3D ESPCN (computationally efficient SR network) Variational dropout using a dual-network architecture

- Intrinsic and Parameter Uncertainty modelled using Heteroscedastic noise model and
- Even a simple baseline model (3 hidden layers) outperforms RF based regression



- Deep Learning IQT:
  - Generalises 3D ESPCN (computationally efficient SR network)
     Intrinsic and Parameter Uncertainty modelled using Heteroscedastic noise model and Variational dropout using a dual-network architecture
     Even a simple baseline model (3 hidden layers) outperforms RF based regression
     Top models quantify Intrinsic + Parameter Uncertainty



#### • Deep Learning IQT:

Generalises 3D ESPCN (computationally efficient SR network) Variational dropout using a dual-network architecture Even a simple baseline model (3 hidden layers) outperforms RF based regression Top models quantify Intrinsic + Parameter Uncertainty Predictive uncertainty flags pathology and can be used as a safeguard

- Intrinsic and Parameter Uncertainty modelled using Heteroscedastic noise model and



#### • Deep Learning IQT:

Generalises 3D ESPCN (computationally efficient SR network) Variational dropout using a dual-network architecture Even a simple baseline model (3 hidden layers) outperforms RF based regression Top models quantify Intrinsic + Parameter Uncertainty Predictive uncertainty flags pathology and can be used as a safeguard

• Perspectives and Challenges:

- Intrinsic and Parameter Uncertainty modelled using Heteroscedastic noise model and



#### • Deep Learning IQT:

Generalises 3D ESPCN (computationally efficient SR network) Variational dropout using a dual-network architecture Even a simple baseline model (3 hidden layers) outperforms RF based regression Top models quantify Intrinsic + Parameter Uncertainty Predictive uncertainty flags pathology and can be used as a safeguard

Perspectives and Challenges: Quantify performance in abnormal tissue (comparison with other methods requires data)

- Intrinsic and Parameter Uncertainty modelled using Heteroscedastic noise model and



#### • Deep Learning IQT:

Generalises 3D ESPCN (computationally efficient SR network) Variational dropout using a dual-network architecture Even a simple baseline model (3 hidden layers) outperforms RF based regression Top models quantify Intrinsic + Parameter Uncertainty Predictive uncertainty flags pathology and can be used as a safeguard

#### Perspectives and Challenges:

- Explore other problems (parameter mapping, data harmonisation, etc...)

- Intrinsic and Parameter Uncertainty modelled using Heteroscedastic noise model and

Quantify performance in abnormal tissue (comparison with other methods requires data)



#### REFERENCES

[Alexander 2014] D. C. Alexander, D Zikic, J Zhang, H Zhang, and A Criminisi, Image Quality Transfer via Random Forest Regression: Applications in Diffusion MRI MICCAI 2014, Part III, LNCS 8675, pp. 225–232, 2014

 [Tanno 2016] R Tanno, A Ghosh, F Grussu, E Kaden, A Criminisi and D C. Alexander Bayesian Image Quality Transfer, MICCAI 2016 pp 265-273, 2016

 [Alexander 2017] D C Alexander, D Zikic, A Ghosh, R Tanno, V Wottschel, J Zhang, E Kaden, T B Dyrby, S N Sotiropoulos, H Zhang, A Criminisi, Image quality transfer and applications in diffusion MRI, NeuroImage Volume 152, 15 May 2017, Pages 283–298

**[Tanno 2017]** R Tanno, D. E. Worrall, A Ghosh, E Kaden, S N. Sotiropoulos, A Criminisi, D. C. Alexander. **Bayesian Image Quality Transfer with CNNs: Exploring Uncertainty in dMRI Super-Resolution**, MICCAI 2017



## THANK YOU









#### https://github.com/ucl-mig/iqt

