

Image quality assessment by overlapping task-specific and task-agnostic measures: application to prostate multiparametric MR images for cancer segmentation

Shaheer U. Saeed

shaheer.saeed.17@ucl.ac.uk

Centre for Medical Image Computing; Wellcome/EPSRC Centre for Interventional & Surgical Sciences; Department of Medical Physics & Biomedical Engineering, University College London, London, UK.

Wen Yan

Centre for Medical Image Computing; Wellcome/EPSRC Centre for Interventional & Surgical Sciences; and Department of Medical Physics & Biomedical Engineering, University College London, London, UK; City University of Hong Kong, Department of Electrical Engineering, Hong Kong, China.

Yunguan Fu

Centre for Medical Image Computing; Wellcome/EPSRC Centre for Interventional & Surgical Sciences; and Department of Medical Physics & Biomedical Engineering, University College London; and InstaDeep, London, UK.

Francesco Giganti

Department of Radiology, University College London Hospital NHS Foundation Trust; and Division of Surgery and Interventional Science, University College London, London, UK.

Qianye Yang

Centre for Medical Image Computing; Wellcome/EPSRC Centre for Interventional & Surgical Sciences; Department of Medical Physics & Biomedical Engineering, University College London, London, UK.

Zachary M. C. Baum

Centre for Medical Image Computing; Wellcome/EPSRC Centre for Interventional & Surgical Sciences; Department of Medical Physics & Biomedical Engineering, University College London, London, UK.

Mirabela Rusu

Department of Radiology, Stanford University, Stanford, California, USA.

Richard E. Fan

Department of Urology, Stanford University, Stanford, California, USA.

Geoffrey A. Sonn

Department of Radiology; and Department of Urology, Stanford University, Stanford, California, USA.

Mark Emberton

Department of Urology, University College Hospital NHS Foundation Trust; and Division of Surgery and Interventional Science, University College London, London, UK.

Dean C. Barratt

Centre for Medical Image Computing; Wellcome/EPSRC Centre for Interventional & Surgical Sciences; Department of Medical Physics & Biomedical Engineering, University College London, London, UK.

Yipeng Hu

Centre for Medical Image Computing; Wellcome/EPSRC Centre for Interventional & Surgical Sciences; Department of Medical Physics & Biomedical Engineering, University College London, London, UK.

Abstract

Image quality assessment (IQA) in medical imaging can be used to ensure that downstream clinical tasks can be reliably performed. Quantifying the impact of an image on the specific target tasks, also named as task amenability, is needed. A *task-specific* IQA has recently been proposed to learn an image-amenability-predicting controller simultaneously with a target task predictor. This allows for the trained IQA controller to measure the impact an image has on the target task performance, when this task is performed using the predictor,

e.g. segmentation and classification neural networks in modern clinical applications. In this work, we propose an extension to this task-specific IQA approach, by adding a *task-agnostic* IQA based on auto-encoding as the target task ¹. Analysing the intersection between low-quality images, deemed by both the task-specific and task-agnostic IQA, may help to differentiate the underpinning factors that caused the poor target task performance. For example, common imaging artefacts may not adversely affect the target task, which would lead to a low task-agnostic quality and a high task-specific quality, whilst individual cases considered clinically challenging, which can not be improved by better imaging equipment or protocols, is likely to result in a high task-agnostic quality but a low task-specific quality. We first describe a flexible reward shaping strategy which allows for the adjustment of weighting between task-agnostic and task-specific quality scoring. Furthermore, we evaluate the proposed reinforcement learning algorithm, using a clinically challenging target task of prostate tumour segmentation on multiparametric magnetic resonance (mpMR) images. Based on experimental results using mpMR images from 850 patients, it was found that *a)* The task-agnostic IQA may identify artefacts, but with limited impact on the accuracy of cancer segmentation networks. A Dice score of 0.367 ± 0.017 was obtained after rejecting 10% of low quality images, compared to 0.354 ± 0.016 from a non-selective baseline; *b)* The task-specific IQA alone improved the performance to 0.415 ± 0.020 , at the same rejection ratio. However, this system indeed rejected both images that impact task performance due to imaging defects and due to being clinically challenging; and *c)* The proposed reward shaping strategy, when the task-agnostic and task-specific IQA are weighted appropriately, successfully identified samples that need re-acquisition due to defected imaging process, as opposed to clinically challenging cases due to low contrast in pathological tissues or other equivocity in radiological presentation.

Our code is available at <https://github.com/s-sd/task-amenability/tree/v1>.

Keywords: Reinforcement learning, Meta-learning, Image quality assessment, Reward shaping, Task amenability

1. Introduction

1.1 Image quality assessment

Image quality assessment (IQA) is utilised extensively in medical imaging and can help to ensure that intended downstream clinical tasks for medical images, such as diagnostic, navigational or therapeutic tasks, can be reliably performed. It has been demonstrated that poor image quality may adversely impact task performance (Davis et al., 2009; Wu et al., 2017; Chow and Paramesran, 2016). In these scenarios IQA can help to ensure that the negatively impacted performance can be counter-acted for example by flagging samples for re-acquisition or defect correction. Most medical images are acquired with an intended clinical task, however, despite the task-dependent nature of medical images, IQA is often studied in task-agnostic settings (Chow and Paramesran, 2016), in which the task is independently performed or, for automated tasks, the task predictors (e.g. machine learning models) are disregarded. Where task-specific IQA, which accounts for the impact of an image on a downstream task, is studied, the impact is usually quantified subjectively by human observers or is learnt from subjective human labels.

1. Although the auto-encoding for image self-reconstruction can be considered a ‘target task’ in the training algorithm (Section 2.2), it is not a clinical task. This target task provides clinical-task-agnostic reward signals for training the task-agnostic IQA.

Both manual and automated methods have been proposed for task-agnostic as well as task-specific IQA. In clinical practice, manually assessing images for their perceived impact on the target task is common practice (Chow and Paramesran, 2016). Usually, manual assessment involves subjectively defined criteria used to assess images where the criteria may or may not account for the impact of image quality on the target task (Chow and Paramesran, 2016; Loizou et al., 2006; Hemmsen et al., 2010; Shima et al., 2007; De Angelis et al., 2007; Chow and Rajagopal, 2015). Automated methods to IQA, either task-specific or task-agnostic, enable reproducible measurements and reduce subjective human interpretation (Chow and Paramesran, 2016). So called no-reference automated methods capture image features common across low or high quality images, in order to perform IQA (Dutta et al., 2013; Kalayeh et al., 2013; Mortamet et al., 2009; Woodard and Carley-Spencer, 2006). The selected common features may be task-specific in some cases, however, these methods are often task-agnostic since the kinds of features selected in general have no bearing on a specific clinical task (e.g. distortion lying outside gland boundaries for gland segmentation). Full- and reduced-reference automated methods use a single or a set of reference images and produce IQA measurements by quantifying similarity to the reference images (Fuderer, 1988; Jiang et al., 2007; Miao et al., 2008; Wang et al., 2004; Kumar and Rattan, 2012; Kumar et al., 2011; Rangaraju et al., 2012; Kowalik-Urbaniak et al., 2014; Huo et al., 2006; Kaufman et al., 1989; Henkelman, 1985; Dietrich et al., 2007; Shiao et al., 2007; Geissler et al., 2007; Salem et al., 2002; Shiao et al., 2007; Choong et al., 2006; Daly, 1992). The selection of the reference, to which similarity is computed, is usually either done manually or using a set of subjectively defined criteria which may or may not account for impact on a clinical task.

Recent deep learning based approaches to IQA also rely on human labels of IQA and offer fast inference (Wu et al., 2017; Zago et al., 2018; Esses et al., 2018; Baum et al., 2021; Liao et al., 2019; Abdi et al., 2017; Lin et al., 2019; Oksuz et al., 2020). Although these methods can provide reproducible and repeatable measurements, the extent to which they can quantify the impact of an image on the target task remains unanswered.

Perhaps more importantly, many downstream clinical target tasks have been automated using machine learning models, and therefore the human-perceived quality, even when they are intended to be task-specific (Wu et al., 2017; Esses et al., 2018; Eck et al., 2015; Racine et al., 2016), may not be representative of the actual impact of an image on the task to be completed by computational models. This is also true for subjective judgement that is involved in other manual or automated no-, reduced- or full-reference methods. Some works do objectively quantify image impact on a downstream task, however, they are specific to particular applications or modalities, for example for IQA or under-sampling in MR images (Mortamet et al., 2009; Woodard and Carley-Spencer, 2006; Razumov et al., 2021), IQA for CT images (Eck et al., 2015; Racine et al., 2016) and synthetic data selection for data augmentation (Ye et al., 2020). In a typical scenario where imaging artefacts lie outside regions of interest for a specific task, general quantification of IQA may not indicate usefulness of the images. For a target task of tumour segmentation on MR images, if modality-specific artefacts, such as motion artefact and magnetic field distortion, are present but of great distance from the gland, then segmentation performance may not be impacted negatively by these artefacts. We present examples of this type in Sect. 4.

1.2 Task-specific image quality: task amenability

In our previous work (Saeed et al., 2021) we proposed to use the term ‘task amenability’ to define the usefulness of an image for a particular target task. This is a task-specific measure of IQA, by which selection of images based on their task amenability may lead to improved target task performance. Task amenability-based selection of images may be useful under several potential scenarios such as for meeting a clinically defined requirement on task performance by removing images with poor task amenability, filtering images in order to re-acquire for cases where the target task cannot be reliably performed due to poor quality, and real-time acquisition quality or operator skill feedback. ‘Task amenability’ is used interchangeably with ‘task-specific image quality’ in this paper henceforth.

In order to quantify task amenability, a controller, which selects task amenable image data, and a task predictor, which performs the target task, are jointly trained. In this formulation the controller decisions are informed directly by the task performance, such that optimising the controller is dependent on the task predictor being optimised - a meta-learning problem. We thus formulated a reinforcement learning (RL) framework for this meta-learning problem, where data weighting/selection is performed in order to optimise target task performance (Saeed et al., 2021).

Modifying meta- or hyper-parameters to optimise task performance using RL algorithms has been proposed previously, including selecting augmentation policies (Cubuk et al., 2019; Zhang et al., 2019), selecting convolutional filters and other neural network hyper-parameters (Zoph and Le, 2017), weighting different imaging modalities (Wang et al., 2020), and valuating or weighting training data (Yoon et al., 2020); the target task may be any task performed by a neural network, for example classification or regression. The independently-proposed task amenability method (Saeed et al., 2021) shares some similarity with the data valuation approach (Yoon et al., 2020), however, different from Yoon et al. (2020), our work investigated reward formulations using controller weighted/ selected data and the use of the controller on holdout data, in addition to other methodological differences.

1.3 Intersection between task-specific and task-agnostic low-quality images

It is worth noting that our previously proposed framework is task-specific and therefore may not be able to distinguish between poor-quality samples due to imaging defects and those due to clinical difficulty. For example, for the diagnostic task of tumour segmentation, despite no visible imaging artefacts, an image may still be considered challenging because of low tissue contrast between the tumour and surrounding regions or because of a very small tumour size. It may have a different outcome compared to cases where imaging defects, such as artefacts or noise, cause reduced task performance. The former cases that are clinically challenging may require further expert input or consensus, whereas latter cases with imaging defects may need to be re-acquired.

Therefore, it may be useful to perform task-agnostic IQA, when imaging defects need to be identified. Task-agnostic methods to IQA quantify image quality as a metric based on the presence of severe artefacts, excessive noise or other general imaging traits for low quality. Those observable general quality issues include unexpected low spatial resolution, electromagnetic noise patterns, motion-caused blurring and distortion due to imperfect reconstruction, which often determine human labels for poor task-agnostic quality. Although

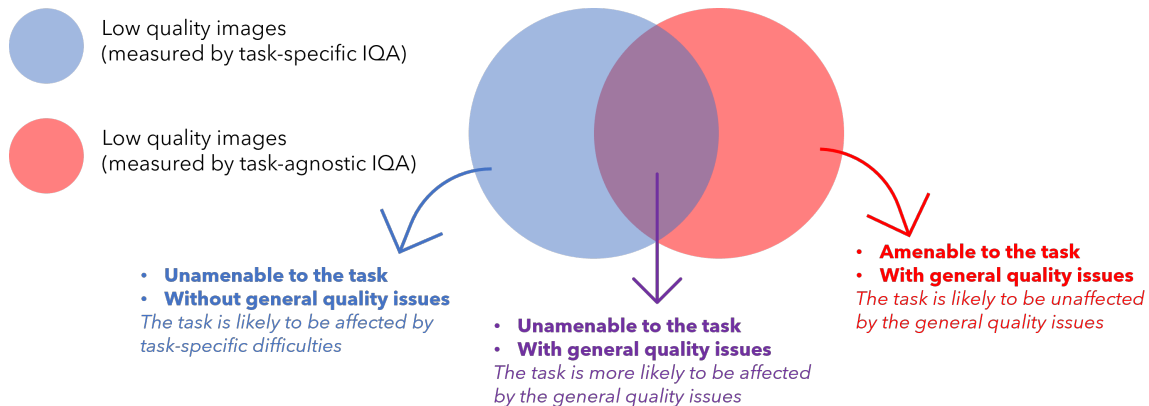


Figure 1: Venn diagram to illustrate the low quality images with respect to task-specific and task-agnostic IQA measures, and their intersections.

these may not affect the target task and/or without necessarily well-understood underlying causes or definitive explanations, these task-agnostic quality issues should be flagged to indicate potentially avoidable or amendable imaging defects. Examples of previous work include those using auto-encoders as feature extractors for clustering or using supervised learning (Li et al., 2020; Yang et al., 2017).

Fig 1 illustrates the relationship between task-specific and task-agnostic qualities discussed in this work as a Venn diagram. Task-specific IQA identifies all samples (blue circle) that negatively impact task performance regardless of the underlying reason. In contrast, task-agnostic IQA may be able to identify images (red circle) with general quality issues, as discussed above, regardless of their impact on the target task. Therefore, image samples that are identified by the task-specific IQA but *not* by the task-agnostic IQA (blue area excluding the purple overlap) are likely to be without readily-identifiable general imaging quality issues, thus more likely to be difficult cases due to patient-specific anatomy and pathology. Whilst, image samples that suffer from general quality issues, such as imaging defects, *and* also negatively impact task performance shall lie in the (purple) overlap between samples identified by both task-specific- and task-agnostic IQA.

In this work, we propose to incorporate task-agnostic IQA by a flexible reward shaping strategy, which allows for the weighting between the two IQA types to be adjusted for the application of interest. To evaluate images based on their task-agnostic image quality, we propose to use auto-encoding as the target task, trained in the same RL algorithm. It is trained with a reward based on image reconstruction error, since random noise and rare artefacts may not be reconstructed well by auto-encoders (Vincent et al., 2008, 2010; Gondara, 2016). Once such a task-agnostic IQA is learned, it can be used to shape the originally-proposed task-specific reward. This allows to train a final IQA controller that ranks the images by both task-specific and task-agnostic measures, with which the lowest-ranked images are much more likely to be those depicted in the overlap area of Fig 1, compared to either of these two IQA approaches alone. A completely task-specific IQA controller and a completely task-agnostic IQA controller are both a special case of the

proposed reward shaping mechanism, by configuring a hyper-parameter that controls this quality-ranking.

1.4 Multiparametric MR quality for prostate cancer segmentation

Prostate cancer is among the most commonly occurring malignancies in the world (Merriel et al., 2018). In current prostate cancer patient care, biopsy-based histopathology outcome remains the diagnostic gold-standard. However, the emergence of multiparametric magnetic resonance (mpMR) imaging does not only offer a potential non-invasive alternative, but also a role in better localising tumour for targeted biopsy or other increasingly localised treatment, such as several focal ablation options (Chen et al., 2020; Ahmed et al., 2011; Marshall and Taneja, 2015; Ahdoot et al., 2019).

Detecting, grading and segmenting tumours from prostate mpMR are known challenging radiological tasks, with a reported 7-14% of missed clinically significant cancers (Ahmed et al., 2017; Rouvière et al., 2019), a high inter-observer variability (Chen et al., 2020) and a strong dependency on the image quality (de Rooij et al., 2020), echoed by ongoing effort in standardising both the radiologist reporting (Weinreb et al., 2016) and the IQA for mpMR (Giganti et al., 2020).

The clinical challenges may directly lead to high-variance in both radiological and histopathological labels, for training machine learning models to aid this diagnostic task. Recently proposed methods (Lavasani et al., 2018; Chen et al., 2021; Cao et al., 2019), mostly based on deep neural networks, reported a relatively low segmentation accuracy in Dice ranging from 20% to 40%. When analysing these machine learning models to improve the target task performance, the two types of image quality issues discussed in Sect. 1.3 have both been observed. Common image quality issues such as grain, susceptibility or distortion (Mazaheri et al., 2013) can adversely impact lesion segmentation, so can a small or ambiguous lesion at early stage of its malignancy. The desirable image quality is also likely to be different for different target tasks, for example gland segmentation (Soerensen et al., 2021; Ghavami et al., 2019). As discussed in previous sections, a re-scan at an experienced cancer centre with a well-tuned mpMR protocol may help certain type of patient cases, but others may require a second radiologist reading or further biopsy.

In this work, we use this clinically-difficult and imaging-dependent target task to test the proposed IQA system, 1) for assessing the mpMR image quality and 2) for further sub-typing the predicted quality issues. As discussed above, both of these abilities are important in informing subsequent clinical decisions, when these machine learning models are deployed in prostate cancer patient care.

1.5 Contribution summary

This paper extends the previously published work (Saeed et al., 2021) and includes original contributions as follows. **First**, 1a) we review the proposed IQA formulation based on task amenability; and 1b) summarise the results based on a previously described interventional application, in which transrectal ultrasound (TRUS) image quality was evaluated with respect to two different target tasks, classifying and segmenting prostate glands; and 1c) we further consolidate the proposed IQA approach with results from new ablation studies to better evaluate different components in the framework; **Second**, 2a) we describe a new

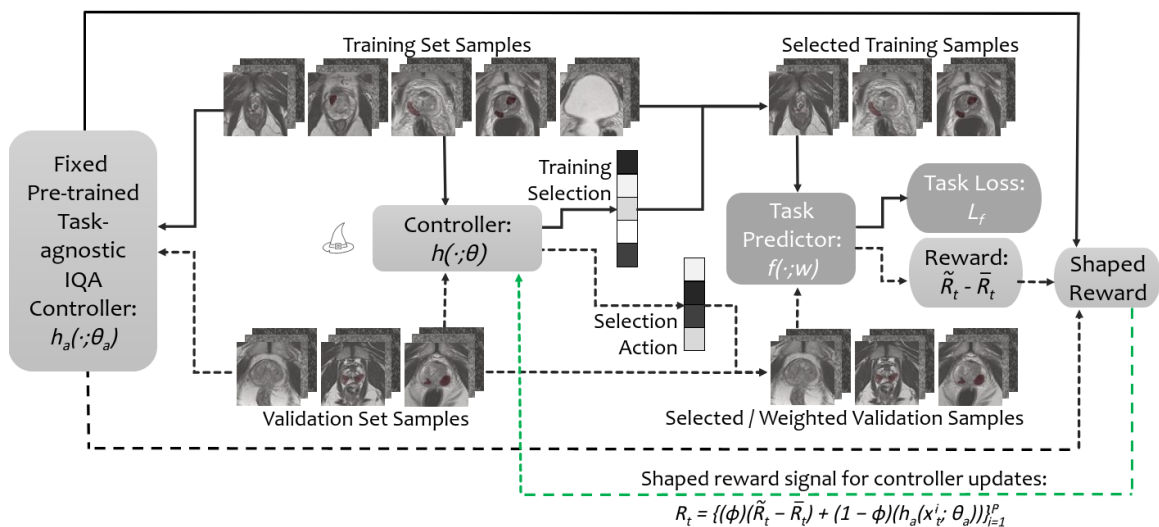


Figure 2: Summary of the IQA framework, with which all types of IQA controllers discussed in this work can be trained. The black solid lines indicate the pathway for the train set data and black dashed lines indicate the pathway for the validation set data; green dashed line indicates reward used to update controller).

extension to use an independently-learned task-agnostic IQA to augment the original task-specific formulation; and 2b) propose a novel flexible reward shaping strategy to learn such a weighted IQA using a single RL-based framework; and 2c) we describe a new set of experimental results, based on a new clinical application, deep-learning-based cancer segmentation from prostate mpMR images, in order to evaluate both the original and the extended IQA approaches.

2. Methods

2.1 Learning image quality assessment

2.1.1 IMAGE QUALITY ASSESSMENT

The proposed image quality assessment framework is comprised of two parametric functions: 1) the task predictor, $f(\cdot; w) : \mathcal{X} \rightarrow \mathcal{Y}$, which performs the target task by producing a prediction $y \in \mathcal{Y}$ for a given image sample $x \in \mathcal{X}$; 2) the controller, $h(\cdot; \theta) : \mathcal{X} \rightarrow [0, 1]$, which scores image samples x based on their task-specific quality². The image and label domains for the target task are denoted by \mathcal{X} and \mathcal{Y} , respectively. Thus in this formulation, we define the image and joint image-label distributions as \mathcal{P}_X and \mathcal{P}_{XY} , respectively. These distributions have probability density functions $p(x)$ and $p(x, y)$, respectively.

The loss function to be minimised for optimising the task predictor is $L_f : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}_{\geq 0}$. This loss function L_f measures how well the target task is performed by the task predictor

2. Task-specific quality means quality scores that represent the impact of an image on a target task; the target task may include any machine learning task including classification, segmentation, or self-reconstruction where these task may not necessarily be clinical tasks.

$f(\cdot; w)$. This loss weighted by controller-predicted IQA scores can be minimised such that poor target task performance for image samples with low controller-predicted scores are weighted less:

$$\min_w \mathbb{E}_{(x,y) \sim \mathcal{P}_{XY}} [L_f(f(x; w), y)h(x; \theta)], \quad (1)$$

On the other hand the metric function $L_h : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}_{\geq 0}$ measures task performance on a validation set. Intuitively, performing the task on images with lower task-specific quality tends to be difficult. To encourage the controller to predict lower quality scores for samples with higher metric values (lower task performance), the following weighted metric function can be minimised:

$$\min_{\theta} \mathbb{E}_{(x,y) \sim \mathcal{P}_{XY}} [L_h(f(x; w), y)h(x; \theta)], \quad (2)$$

$$\text{s.t. } \mathbb{E}_{x \sim \mathcal{P}_X} [h(x; \theta)] \geq c > 0 \quad (3)$$

Here, c is a constant that ensures non-zero quality scores.

The controller thus learns to predict task-specific quality scores for image samples while the task predictor learns to perform the target task. The IQA framework can thus be assembled as the following bi-level minimisation problem (Sinha et al., 2018):

$$\min_{\theta} \mathbb{E}_{(x,y) \sim \mathcal{P}_{XY}} [L_h(f(x; w^*), y)h(x; \theta)], \quad (4a)$$

$$\text{s.t. } w^* = \arg \min_w \mathbb{E}_{(x,y) \sim \mathcal{P}_{XY}} [L_f(f(x; w), y)h(x; \theta)], \quad (4b)$$

$$\mathbb{E}_{x \sim \mathcal{P}_X} [h(x; \theta)] \geq c > 0. \quad (4c)$$

A re-formulation to permit sample selection based on task-specific quality scores, where the the data x and (x, y) are sampled from the controller-selected or -sampled distributions \mathcal{P}_X^h and \mathcal{P}_{XY}^h , with probability density functions $p^h(x) \propto p(x)h(x; \theta)$ and $p^h(x, y) \propto p(x, y)h(x; \theta)$ respectively, can be defined as follows:

$$\min_{\theta} \mathbb{E}_{(x,y) \sim \mathcal{P}_{XY}^h} [L_h(f(x; w^*), y)], \quad (5a)$$

$$\text{s.t. } w^* = \arg \min_w \mathbb{E}_{(x,y) \sim \mathcal{P}_{XY}^h} [L_f(f(x; w), y)], \quad (5b)$$

$$\mathbb{E}_{x \sim \mathcal{P}_X^h} [1] \geq c > 0. \quad (5c)$$

2.1.2 REINFORCEMENT LEARNING

This bi-level minimisation problem can be formulated in a RL setting. We propose to formulate this problem in a RL-based meta learning setting where the controller interacting with the task predictor via sample-selection or -weighting can be considered a Markov decision process (MDP). The MDP in this problem setting can be described by a 5-tuple $(\mathcal{S}, \mathcal{A}, p, r, \pi)$. The state transition distribution $p : \mathcal{S} \times \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$ denotes the probability of the next state s_{t+1} given the current state s_t and action a_t ; here s_{t+1} , $s_t \in \mathcal{S}$ and $a_t \in \mathcal{A}$, where \mathcal{S} is the state space and \mathcal{A} is the action space. The reward function is $r : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ and $R_t = r(s_t, a_t)$ denotes the reward given s_t and a_t . The policy is $\pi(a_t | s_t) : \mathcal{S} \times \mathcal{A} \in [0, 1]$, which represents probability of performing action a_t given s_t . The rewards accumulated

starting from time-step t can be denoted by: $Q^\pi(s_t, a_t) = \sum_{k=0}^T \gamma^k R_{t+k}$, where the discount factor for future rewards is $\gamma \in [0, 1]$. A trajectory or sequence can be observed as the MDP interaction takes place, the sequence takes the form $(s_1, a_1, R_1, s_2, a_2, R_2, \dots, s_T, a_T, R_T)$. If the interactions take place according to a parameterised policy π_θ , then the objective in RL is to learn optimal policy parameters $\theta^* = \operatorname{argmax}_\theta \mathbb{E}_{\pi_\theta} [Q^\pi(s_t, a_t)]$.

2.1.3 PREDICTING IMAGE QUALITY USING REINFORCEMENT LEARNING

Formulating the IQA framework, outlined in Sect. 2.1, as a RL-based meta-learning problem, the finite dataset together with the task predictor can be considered to be contained inside an environment with which the controller interacts. Then at time-step t , from the training dataset $\mathcal{D}_{\text{train}} = \{(x_t^i, y_t^i)\}_{i=1}^N$, a batch of samples $\mathcal{B}_t = \{(x_t^i, y_t^i)\}_{i=1}^B$ together with the task predictor $f(\cdot; w_t)$ may be considered an observed state $s_t = (f(\cdot; w_t), \mathcal{B}_t)$. The controller $h(\cdot; \theta)$ outputs sampling probabilities which dictate selection decisions for samples based on $a_t^i \sim \text{Bernoulli}(h(x_t^i; \theta))$ where action $a_t = \{a_t^i\}_{i=1}^B \in \{0, 1\}^B$ contains the selection decision for samples in the batch. Since the actions are binary selection decisions, at each time-step t , the possible actions are 2^B . If $a_t^i = 1$, sample (x_t^i, y_t^i) is selected for training the task predictor. The policy $\pi_\theta(a_t | s_t)$ may be defined as $\log \pi_\theta(a_t | s_t) = \sum_{i=1}^D h(x_t^i; \theta) a_t^i + (1 - h(x_t^i; \theta))(1 - a_t^i)$.

In this RL-based meta-learning framework, the reward is computed based on the task predictor’s performance $\{l_t^j\}_{j=1}^M = \{L_h(f(x^j; w_t), y^j)\}_{j=1}^M$ on a validation set $\mathcal{D}_{\text{val}} = \{(x^j, y^j)\}_{j=1}^M$. The controller outputs for the validation set $\{h^j\}_{j=1}^M = \{h(x^j; \theta)\}_{j=1}^M$ may be utilised to formulate the final reward. While several strategies exist for reward computation, we investigate three strategies to compute an un-clipped reward \tilde{R}_t . Three strategies to compute \tilde{R}_t are as follows:

1. $\tilde{R}_{\text{avg},t} = -\frac{1}{M} \sum_{j=1}^M l_t^j$, the average performance,
2. $\tilde{R}_{\text{w},t} = -\frac{1}{M} \sum_{j=1}^M l_t^j h^j$, the weighted sum,
3. $\tilde{R}_{\text{sel},t} = -\frac{1}{M'} \sum_{j'=1}^{M'} l_t^{j'}$, the average of the selected M' samples;

For the selective reward formulation $\tilde{R}_{\text{sel},t} \{j'\}_{j'=1}^{M'} \subseteq \{j\}_{j=1}^M$ with $h^{j'} \leq h^{k'}$ for $\forall k' \in \{j'\}^c$, $\forall j' \in \{j'\}$, i.e. the unclipped reward $\tilde{R}_{\text{sel},t}$ is the average of $\{l_{j'}\}$ from the subset of $M' = \lfloor (1 - s^{rej})M \rfloor$ samples, by removing the first $s^{rej} \times 100\%$ samples, after sorting h_j in decreasing order. The first reward formulation $\tilde{R}_{\text{avg},t}$ requires pre-selection of data with high task-specific quality whereas the other two reward formulations do not require any pre-selection of data and the validation set may be of mixed quality. In the selective and weighted reward formulations data is selected or weighted based on controller-predicted task-specific quality. The use of a validation set ensures that the system encourages generalisability therefore extreme cases, where the controller values one sample very highly and the remaining samples as very low, are discouraged. The validation set is formed of data which is weighted or selected based on task-specific quality, either using on human labels or controller predictions. This ensures that selection based on task-specific quality, in the train set, is encouraged as opposed to selecting all samples to improve generalisability. It should be noted that other strategies may be used to compute \tilde{R}_t , however, in this work we only evaluate the three outlined above.

To form the final reward R_t , we clip the performance measure \tilde{R}_t using a clipping quantity \bar{R}_t . The final reward thus takes the form:

$$R_t = \tilde{R}_t - \bar{R}_t \quad (6)$$

Similar to \tilde{R}_t , several different formulations may be used for \bar{R}_t . In this work we use a moving average $\bar{R}_t = \alpha_R \bar{R}_{t-1} + (1 - \alpha_R) \tilde{R}_t$, where α_R is a hyper-parameter set to 0.9. It should be noted that, although optional, moving average-based clipping serves to promote continual improvement. A random selection baseline model or a non-selective baseline model may also be used as \bar{R}_t , however, these are not investigated in this work.

The RL-based meta-learning framework to learn IQA is summarised in Fig. 2 and is outlined in Algo. 1.

2.2 Learning task-agnostic image quality assessment

As outlined in Sect. 1, under certain circumstances, it may be useful to learn task-agnostic IQA. For example when imaging protocols may need to be fine-tuned in order to remove noise and artefacts or when a particular target task may not be known. It is possible to learn such a task-agnostic IQA in the framework presented in Sect. 2.1.3 using auto-encoding as the target task. It should be noted that while auto-encoding or self-reconstruction is a target task in this RL framework, it is not a clinical task and is only used for the purpose of learning a task-agnostic IQA.

For the auto-encoding target task, the label distribution is set as the image distribution $\mathcal{Y} = \mathcal{X}$. With the L_h based on image reconstruction error, such as mean squared error $1/n \sum_{i=1}^n (y_t^i - \hat{y}_t^i)^2$ where \hat{y} is the task predictor-predicted label, in the presented framework. Features not common across the entire distribution \mathcal{X} such as random noise and randomly placed artefacts may be difficult to reconstruct (Vincent et al., 2008, 2010; Gondara, 2016). The intuition is that due to a higher reconstruction error for such samples, these samples are to be valued lower by the controller. In addition to detecting samples with random noise and artefacts, this scheme may also be used for unsupervised anomaly detection although this is not discussed further in this work.

2.3 Shaping task-specific rewards to weight task-agnostic measure

Let $h_a(\cdot; \theta_a)$ be a pre-trained task-agnostic quality-predicting controller, which has been trained as described in Sect. 2.2. This fixed pre-trained task-agnostic quality-predicting controller is used to shape the task-specific rewards in order to learn a weighted or overlapping measure of image quality. A new controller can then be trained with a shaped reward signal where the weighting between the task-agnostic quality and the task-specific quality can be manually adjusted. We use the term ‘shaped’ for this new reward since, as opposed to the reward formulation in Sect. 2.1.3 with a single reward per-batch, the shaped formulation has per-sample rewards as outlined in Eq. 7. A controller trained using the non-shaped reward may not be able to distinguish between samples that negatively impact task performance due to general quality defects, and samples that negatively impact performance due to clinical difficulty. The learned ability, using the shaped reward signal, to identify only samples that negatively impact task performance due to imaging artefacts or general quality defects may be clinically useful. This ability to distinguish such cases

can help to identify samples with general quality defects that may need re-acquisition. As opposed to simply computing reward using $R_t = \tilde{R}_t - \bar{R}_t$, the shaped reward at time-step t can be computed as follows:

$$R_t = \{\phi(\tilde{R}_t - \bar{R}_t) + (1 - \phi)(h_a(x_t^i; \theta_a))\}_{i=1}^P \quad (7)$$

Here $P = B + M$ is a set of samples formed by the the concatenation of the mini-batch of train samples (with B samples) and the M validation set samples. The per-sample reward for these P samples is computed using Eq. 7. The clipping, in this shaped reward, is only applied to \tilde{R}_t (the task performance measure) and not to the entire reward since the task-agnostic IQA controller h_a is fixed and its predictions are deterministic during training with the shaped reward. This weighted sum of the clipped task performance measure, for the task in question, and the task-agnostic IQA, allows for manual adjustment of the relative importance of the task-agnostic quality and task-specific quality. It should be noted that instead of using h_a in this shaping strategy it is also possible to use human labels of IQA or a different task-specific quality-predicting controller which is specific to a different target task. After training with the shaped reward, the trained controller may be denoted as $h(\cdot; \theta)$. When $\phi = 1$, the shaped reward simplifies to the non-shaped reward $R_t = \tilde{R}_t - \bar{R}_t$, which was introduced in Sect. 2.1.3, thus a fully task-specific IQA is learnt. When $\phi = 0$, the shaped reward simplifies to $R_t = \{h_a(x_t^i; \theta_a)\}_{i=1}^B$ which means that the trained controller will be approximately equal to the pre-trained task-agnostic quality-predicting controller $h(\cdot; \theta) \approx h_a(\cdot; \theta_a)$. In this work, for notational convenience, wherever we use $\phi = 0$, we report results for $h_a(\cdot; \theta_a)$ directly and do not train $h(\cdot; \theta)$.

Algorithm 1: Image quality assessment

Data: Training dataset $\mathcal{D}_{\text{train}}$ and validation dataset \mathcal{D}_{val} .

Result: Task predictor $f(\cdot; w)$ and controller $h(\cdot; \theta)$.

while *not converged* **do**

for $k \leftarrow 1$ **to** K **do**

for $t \leftarrow 1$ **to** T **do**

 Sample a mini-batch $\mathcal{B}_t = \{(x_t^i, y_t^i)\}_{i=1}^B$ from $\mathcal{D}_{\text{train}}$;

 Compute selection probabilities $\{h(x_t^i; \theta_t)\}_{i=1}^B$;

 Sample actions $a_t = \{a_t^i\}_{i=1}^B$ w.r.t. $a_t^i \sim \text{Bernoulli}(h(x_t^i; \theta))$;

 Selected samples $\mathcal{B}_{t,\text{selected}}$ from \mathcal{B}_t ;

 Update predictor $f(\cdot; w_t)$ using $\mathcal{B}_{t,\text{selected}}$;

 Compute reward R_t ;

end

 Collect one episode $\{\mathcal{B}_t, a_t, R_t\}_{t=1}^T$;

end

 Update controller $h(\cdot; \theta)$ using reinforcement learning algorithm;

end

3. Experiments

3.1 Ultrasound image quality for prostate detection and segmentation

3.1.1 PROSTATE CLASSIFICATION AND SEGMENTATION NETWORKS AND THE CONTROLLER ARCHITECTURE

For the prostate presence classification task, which is a binary classification of whether a 2D US slice contains the prostate gland or not, the Alex-Net (Krizhevsky et al., 2012) was used as the task predictor with a cross-entropy loss function and reward based on classification accuracy (Acc.), i.e. classification correction rate. The controller was trained using the deep deterministic policy gradient (DDPG) RL algorithm (Lillicrap et al., 2019) where both the actor and critic used a 3-layer convolutional encoder for the image before passing to 3 fully connected layers.

For the prostate gland segmentation task, which is a segmentation of the gland on a 2D slice of US, a 2D U-Net (Ronneberger et al., 2015) was used as the task predictor with a pixel-wise cross-entropy loss and reward based on mean binary Dice score (Dice). The controller training and architecture details are the same as the prostate presence classification task.

Further implementation details can be found in the GitHub repository: <https://github.com/s-sd/task-amenability/tree/v1>.

3.1.2 EVALUATING THE TASK AMENABILITY AGENT

The experiments for the two tasks of prostate presence classification and gland segmentation have been previously described in Saeed et al. (2021) and the results are summarised in Sect. 4. For these experiments, the task-specific IQA was investigated, i.e. the agent was trained with $\phi = 1$. Evaluation of the proposed task amenability agent for both tasks, including the three proposed reward strategies, and comparisons to human labels of IQA were presented in our previous work Saeed et al. (2021).

Acc. and Dice were used as measures of performance for the classification and segmentation tasks, respectively. These serve as direct measures of performance for the task in question and as indirect measures of performance for the controller with respect to the learnt IQA. Standard deviation (St.D.) is reported as a measure of inter-patient variance and T-test results are reported at a significance level of $\alpha = 0.05$, where comparisons are made. Details on controller selection are as described in Sect. 3.2.2.

3.1.3 EXPERIMENTAL DATA

The data was acquired as part of the SmartTarget:Biopsy and SmartTarget:Therapy clinical trials (NCT02290561, NCT02341677). Images were acquired from 259 patients who underwent ultrasound-guided prostate biopsy. In total 50-120 2D frames of TRUS were acquired for each patient using the side-firing transducer of a bi-plane transperineal ultrasound probe (C41L47RP, HI-VISION Preirus, Hitachi Medical Systems Europe). These 2D frames were acquired during manual positioning of a digital transperineal stepper (D&K Technologies GmbH, Barum, Germany) for navigation or the rotation of the stepper with recorded relative angles for scanning the entire gland. The TRUS images were sampled at approximately every 4 degrees resulting in a total of 6712 images. These 2D images were then segmented by three trained biomedical engineering researchers. For the first task of

prostate presence classification, a binary scalar indicating prostate presence was derived from the segmentation for all three observers and then a majority vote was conducted in order to form the final labels for used for this task. For the prostate gland segmentation, binary segmentation masks of the prostate gland were used and a majority vote at the pixel level was used to form the final labels used for this task. The images were split into three sets, train, validation and holdout, with 4689, 1023, and 1000 images from 178, 43, and 38 subjects, respectively. Samples from the TRUS data are shown in Fig. 3.

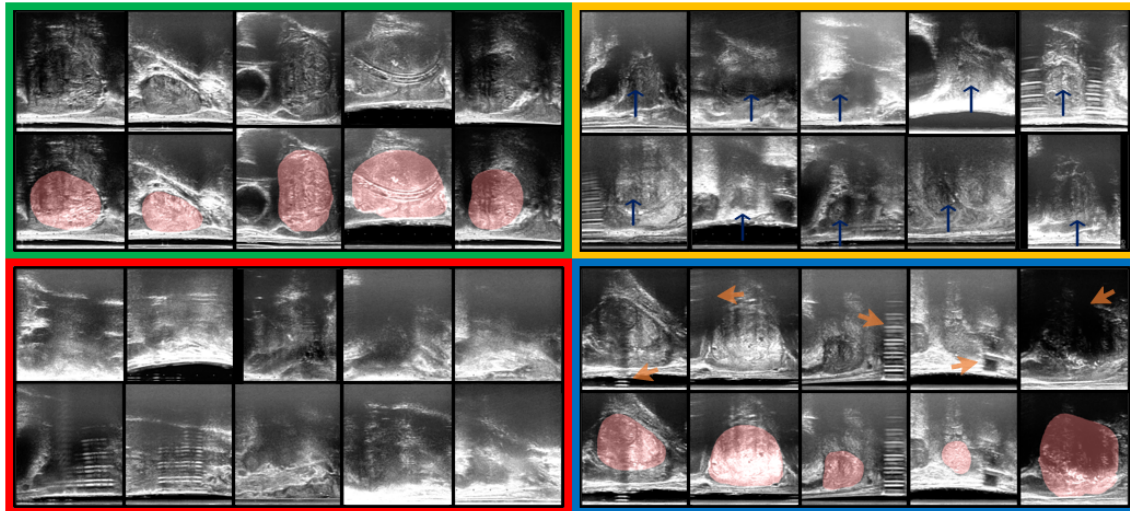


Figure 3: Examples of ultrasound images used in this study. **Top-left** (green): task-amenable images that contain prostate gland (shaded in red); **Bottom-left** (red): images with poor task amenability where recognising the prostate (for classification) and delineating its boundary (for segmentation) is difficult; **Top-right** (yellow), images that likely contain prostate glands (blue arrows) but identifying the complete gland boundaries for segmentation is challenging; and **Bottom-right** (blue): images that contain visible noise and artefacts (orange arrows), but may be amenable to both classification and segmentation tasks.

3.2 Multiparametric MR image quality for tumour segmentation

3.2.1 PROSTATE TUMOUR SEGMENTATION NETWORK AND THE CONTROLLER ARCHITECTURE

A 3D U-Net (Özgün Çiçek et al., 2016) was used as the task predictor with a loss formed of equally weighted pixel-wise cross-entropy and Dice (i.e. $1 - \text{Dice}$) losses and the reward was based on Dice. The reward formulation used for the computation of \tilde{R} was the weighted formulation. A depth of 4 was used with the U-Net where 4 down-sampling and 4 up-sampling layers were used. Each of the down-sampling modules consisted of two convolutional layers with batch normalisation and ReLU activation, and a max-pooling operation. Analogously, de-convolutional layers are used instead in the up-sampling part. Skip connections were used to connect layers in encoding part to the corresponding layers in decoding part.

The input to the network was in the form of a 3-channel 3D image with the channels corresponding to T2-weighted, diffusion-weighted with high b-value and apparent diffusion coefficient MR images. The output is a one-channel probability map of the lesion for computing the loss, which can then be converted to a binary map using thresholding when computing testing results. Other network training details are described in Sect. 3.2.2.

For experiments with the prostate mpMR images, the deep deterministic policy gradient (DDPG) algorithm (Lillicrap et al., 2019) was used as the RL algorithm for training with the only difference compared to experiments using the TRUS data being that the 3-layer convolutional encoders for the actor and critic networks used 3D convolutions rather than 2D. This agent may be considered as having $\phi = 1$.

More implementation details can be found in the GitHub repository: <https://github.com/s-sd/task-amenability/tree/v1>.

3.2.2 EVALUATING THE TASK AMENABILITY AGENT

Dice was used a direct performance measure to evaluate the segmentation task and as an indirect measure of performance for the controller with respect to learnt IQA. Standard deviation (St.D.) is reported as a measure of inter-patient variance. Wherever comparisons are made, T-test results with a significance level of 0.05 are reported. Where controller selection is applicable on the holdout set, samples are ordered according to their controller predicted values and the lowest $k \times 100\%$ samples are removed, where k is referred to as the holdout set rejection ratio. This rejection ratio is specified where appropriate. The results are presented in Sect. 4.

3.2.3 EVALUATING THE TASK-AGNOSTIC IQA AGENT

With the prostate mpMR images, we compare the proposed task-agnostic IQA strategy, presented in Sect. 2.2, with task-specific IQA learnt using the framework outlined in Sect. 2.1.3. To train the task-agnostic IQA agent, the actor and critic networks remain the same as in Sect. 3.2.2. The task predictor, is a fully convolutional auto-encoder with 4 down-sampling and 4 up-sampling layers, with convolutions being in 3D. The loss was based on mean squared error (ie. $1/n \sum_{i=1}^n (y_t^i - \hat{y}_t^i)^2$) and the reward was based on mean absolute error (ie. $1/n \sum_{i=1}^n |y_t^i - \hat{y}_t^i|$) which serves as a measure of image reconstruction performance. This agent may be considered as having $\phi = 0$.

3.2.4 EVALUATING THE REWARD-SHAPING IQA NETWORKS

We also evaluate the effect of different values of the weighting between the task-agnostic and task-specific IQA, ϕ , when using the reward shaping strategy presented in Sect. 2.3. Moreover, samples are presented to further qualitatively evaluate the learnt IQA. These results are presented in Sect. 4.

3.2.5 EXPERIMENTAL DATA

There were 878 sets of mpMR image data acquired from 850 prostate cancer patients. Patients data were acquired as part several clinical trials carried out at University College London Hospitals, with a mixture of biopsy and therapy patient cohorts, including

SmartTarget (Hamid et al., 2019), PICTURE (Simmons et al., 2018), ProRAFT (Orczyk et al., 2021), Index (Dickinson et al., 2013), PROMIS (Bosaily et al., 2015) and PROGENY (Linch et al., 2017). All trial patients gave written consents and the ethics was approved as part of the respective trial protocols. Radiologist contours were done manually and were included for all lesions with a likert score equal or greater than three as the radiological ground-truth in this study. All the data volumes were resampled to $0.5 \times 0.5 \times 0.5mm^3$ in 3D space. A region of interest (ROI) with a volume size of $192 \times 192 \times 96$ voxels centring at the prostate was cropped according to the prostate segmentation mask. All the volumes were then normalised to an intensity range of $[0,1]$. The 3D image sequences used in our study include those with T2-weighted, diffusion-weighted with b values of 1000 or 2000, and apparent diffusion coefficient MR, all available. These were then split into three sets, the train, validation and holdout sets with 445, 64 and 128 images in each set, respectively, where each image corresponds to a separate patient.

4. Results

4.1 Evaluation of the task amenability agent

4.1.1 PROSTATE PRESENCE CLASSIFICATION AND GLAND SEGMENTATION FROM TRANS-RECTAL ULTRASOUND DATA

Evaluating the three reward strategies The target task performance for the prostate presence classification and gland segmentation tasks, on controller selected holdout set samples was investigated in Saeed et al. (2021). The system took approximately 12h to train on a single Nvidia Quadro P5000 GPU. For both tasks, for all three tested reward strategies, significantly higher performance was observed compared to the non-selective baseline (p-value<0.001 for all). Comparing the weighted reward formulation to the fixed clean validation set reward formulation, for the classification and segmentation tasks, no statistical significance was found (p-value=0.06 and 0.49 respectively). Contrastingly, comparing the selective reward formulation with the fixed clean reward formulation, statistical significance was observed, with the selective formulation showing inferior performance for both tasks (p-value<0.001 for all). The plots of task performance against rejection ratio are presented in Fig. 5b, for both tasks. In addition to what was reported in Saeed et al. (2021), we include a non-selective baseline in these plots for comparison. The peak classification Acc. are 0.935, 0.932 and 0.913 at 5%, 10% and 5% rejection ratios, for the fixed-, weighted- and selective reward formulations, respectively, while the peak segmentation Dice are 0.891, 0.893 and 0.866 at 20%, 15% and 20% rejection ratios, respectively.

Investigating the impact of the validation set rejection ratio hyper-parameter for the selective validation set strategy In a first ablation study, we investigate the impact of the validation set rejection ratio s^{rej} , when using the selective validation set reward formulation with a holdout set rejection ratio of 0.15, on the learnt IQA using the prostate gland segmentation task. Increasing s^{rej} from 0.00 to 0.30 in increments of 0.05, we observed a statistically significant improvement in performance for each step increase (p-value<0.01 for all) up to $s^{rej} = 0.20$. Overall, performance (Dice) was improved from 0.827 ± 0.011 at $s^{rej} = 0.00$ to 0.882 ± 0.017 at $s^{rej} = 0.20$. Further increasing to a value of $s^{rej} = 0.25$ and comparing with preceding value of $s^{rej} = 0.20$ led to no significance

being found (p-value=0.37). Another step increase, $s^{rej} = 0.30$, led to lower performance compared to preceding value of $s^{rej} = 0.25$, with statistical significance (p-value<0.01).

Investigating the impact of changing RL algorithms In a second ablation study, we investigate the impact of the RL algorithm on the prostate presence classification task. For the weighted reward formulation with a holdout set rejection ratio of 0.05 we observed task performance (Acc.) of 0.929 ± 0.013 using the proximal policy optimisation (PPO) RL algorithm Schulman et al. (2017) and no significance was found when comparing this to the DDPG algorithm which had a performance of 0.926 ± 0.012 (p-value=0.09).

Comparing controller-predicted IQA with human IQA In addition to the plots of task performance against holdout set rejection ratio presented in Fig. 5b, we also present contingency tables comparing controller-learned task amenability with human labels of IQA in Fig. 5a, for the prostate presence classification and gland segmentation tasks. As reported in Saeed et al. (2021), we summarise these results here since they offer interesting insights into the learnt IQA compared to human-judged IQA. For the purpose of comparison, 0.05 and 0.15 of the holdout set are considered to have low controller-predicted values, for the classification and segmentation tasks, respectively. For these comparisons, Cohen’s kappa values of 0.75, 0.51 and 0.30 were found in the classification task, for the fixed clean, weighted- and selective reward formulations, respectively, and with respective kappa values of 0.63, 0.48 and 0.37 obtained in the segmentation task.

Table 1: Results on the controller-selected holdout set (holdout set rejection ratio of 0.10 used).

Task	Reward computation strategy	Mean \pm St.D.
Lesion segmentation (Dice)	Non-selective baseline	0.354 ± 0.016
	$\bar{R}_{w,t}$, weighted validation set, shaped reward ($\phi = 0.00$)	0.367 ± 0.017
	$\bar{R}_{w,t}$, weighted validation set, shaped reward ($\phi = 0.80$)	0.375 ± 0.016
	$\bar{R}_{w,t}$, weighted validation set, shaped reward ($\phi = 0.85$)	0.380 ± 0.021
	$\bar{R}_{w,t}$, weighted validation set, shaped reward ($\phi = 0.90$)	0.388 ± 0.022
	$\bar{R}_{w,t}$, weighted validation set, shaped reward ($\phi = 0.95$)	0.366 ± 0.018
	$\bar{R}_{w,t}$, weighted validation set, shaped reward ($\phi = 0.99$)	0.405 ± 0.019
	$\bar{R}_{w,t}$, weighted validation set, shaped reward ($\phi = 1.00$)	0.415 ± 0.020

Table 2: Dice results for varying ϕ values against varying values of k .

		ϕ			
		0.00	0.85	0.95	1.00
k	0.00	0.354 ± 0.019	0.368 ± 0.020	0.377 ± 0.023	0.375 ± 0.017
	0.05	0.358 ± 0.018	0.381 ± 0.016	0.383 ± 0.019	0.397 ± 0.019
	0.10	0.367 ± 0.017	0.380 ± 0.021	0.396 ± 0.018	0.415 ± 0.020
	0.15	0.370 ± 0.019	0.383 ± 0.018	0.398 ± 0.016	0.408 ± 0.019
	0.20	0.368 ± 0.021	0.376 ± 0.016	0.394 ± 0.020	0.401 ± 0.015
	0.25	0.366 ± 0.017	0.375 ± 0.019	0.392 ± 0.017	0.404 ± 0.018

4.1.2 PROSTATE LESION SEGMENTATION FROM MPMR IMAGES

Evaluating a fully task-specific reward The task performance results for the prostate lesion segmentation task, for controller-selected holdout samples, are presented in Table 1. The approximate training time for the IQA system was 24h on a single Nvidia Tesla V100 GPU. The results for the task-amenability agent, where a fully task-specific IQA is learnt, are those with $\phi = 1$. For this agent, we observed higher performance for the controller-selected holdout set compared with the non-selective baseline ($p < 0.01$). The plot of performance in terms of Dice against holdout set rejection ratio for this agent, presented in Fig. 6b, shows an initial rise followed by a plateau, as opposed to a small decrease after the initial rise which was observed for the prostate presence classification and gland segmentation tasks using the TRUS data. Samples of task predictor-predicted labels for the lesion segmentation task along with ground truth values are presented in Fig. 4.

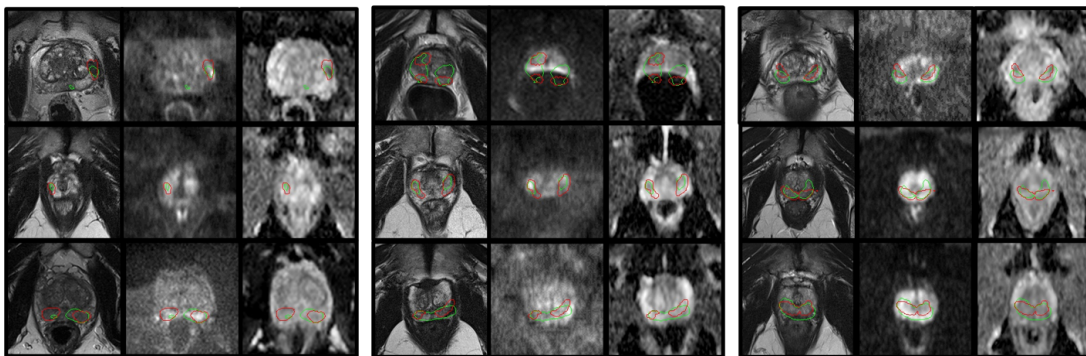


Figure 4: Samples for the prostate lesion segmentation task with the three sub-columns in each column of samples being the three channels T2-weighted, diffusion-weighted, and apparent diffusion coefficient MR. Red is the ground truth and green is the predictor-predicted segmentation.

4.2 Evaluation of the task-agnostic IQA agent

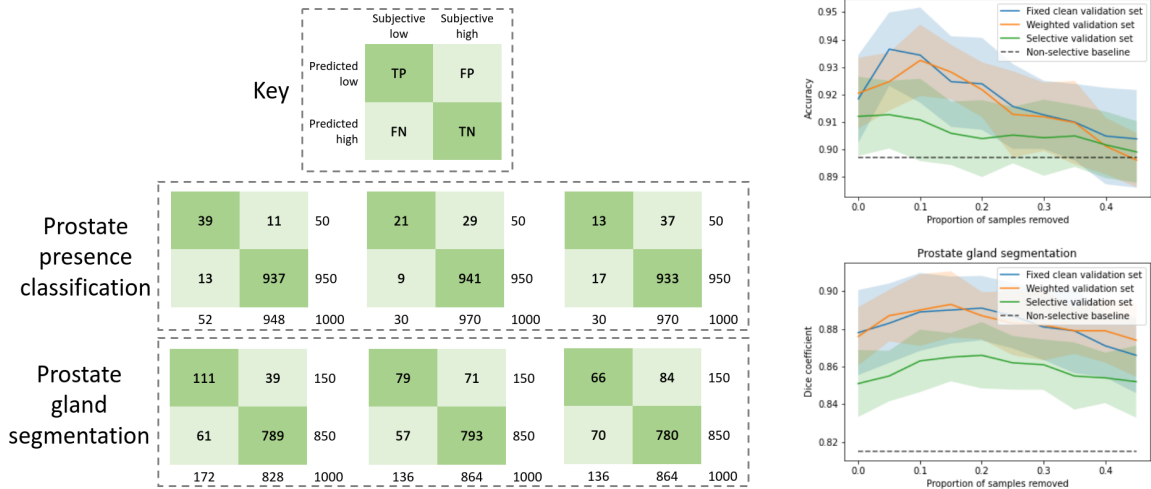
4.2.1 PROSTATE LESION SEGMENTATION FROM MPMR IMAGES

The results for the task-agnostic IQA agent, for the lesion segmentation task using mpMR images, are summarised in Table 1 and Fig. 6. The task-agnostic IQA agent refers to the agent trained with $\phi = 0$. Controller selection of holdout samples for this agent shows a small yet statistically significant improvement in performance compared to the non-selective baseline ($p < 0.01$). A contingency table comparing task-agnostic IQA ($\phi = 0.00$) with task-specific IQA ($\phi = 1.00$) shows the level of disagreement between the two, with a Cohen’s kappa value less than 0.001. Samples for controller predictions are presented in Fig. 7.

4.3 Evaluation of the reward shaping strategies

4.3.1 PROSTATE LESION SEGMENTATION FROM MPMR IMAGES

The results for the shaped reward formulations, for the lesion segmentation task using mpMR images, are summarised in Table 2 and Fig. 6b. Higher performance was observed



(a) Contingency tables comparing subjective labels (high or low task-specific quality labelled by an expert) to controller predictions (high or low task-specific quality determined by the controller using holdout set rejection ratios of 0.05 and 0.15 for the classification and segmentation, respectively) for the different reward formulations, from left to right: 1) fixed clean; 2) weighted; and 3) selective, validation sets. For comparison with controller-predicted “ground-truth” quality, number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) is summarised in the tables. For example, FN is the number of samples predicted as having high controller-predicted quality but low human observer-predicted quality.

(b) Plots of the task performance (in respective Acc. and Dice metrics) against the holdout set rejection ratio for the two tasks of prostate presence classification and gland segmentation.

Figure 5: Summarised results for the prostate presence classification and prostate gland segmentation tasks for the TRUS dataset.

for the controller-selected holdout set compared with the non-selective baseline, for all tested values of ϕ , where the differences were statistically significant ($p < 0.01$ for all). Moreover, the shaped reward formulation with $\phi \geq 0.85$ showed improved performance compared to the shaped reward with $\phi = 0.00$ (task-agnostic IQA agent), with statistical significance ($p < 0.01$ for all). Interestingly, comparing the the formulation with $\phi = 0.00$ to the formulation with $\phi = 0.80$, statistical significance was not found (p -value=0.051).

To further qualitatively assess the proposed IQA agent in identifying the general task-agnostic quality issues, 20 example cases were shown blindly to an experienced radiologist. Among the 8 cases reported as high task-agnostic quality by the IQA agent, the radiologist agreed with 7, with one having “minor artefact”. Among the other 12 cases deemed low task-agnostic quality with varying task-specific quality by the IQA agent, the radiologist agreed with 3 having “significant quality issues that may affect diagnosis” and 6 having “minor quality issues that are unlikely to affect the diagnostic task”, and disagreed with the other 3 having “little quality issues”.

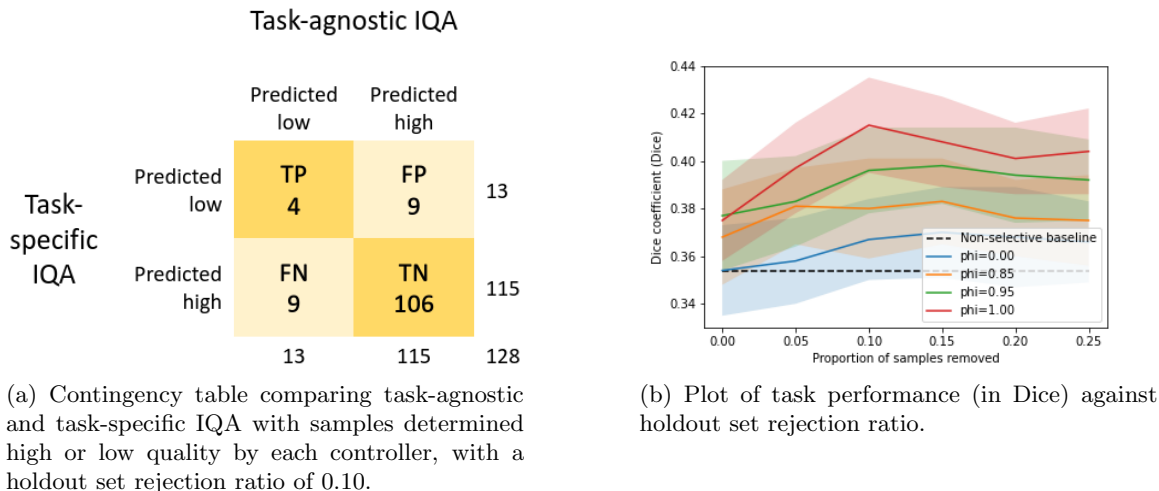


Figure 6: Summarised results for the prostate lesion segmentation task.

5. Discussion and Conclusion

Results from the experiments investigating the three proposed reward strategies, summarised in Fig. 5, show that the selective reward formulation achieved inferior performance compared to the weighted and fixed clean validation set reward strategy. When tuning the s^{rej} parameter for this selective formulation, however, we see a performance increase. Thus, further tuning of this parameter may be required in order to achieve performance comparable to the other reward formulations. Moreover, the selective formulation also offers a mechanism to specify a desirable validation set rejection ratio which may be useful for applications where there is a significant class imbalance problem. Additionally, the selective and weighted reward formulations allow for task-specific IQA to be learnt without any human labels of IQA and the weighted reward formulation does so without any significant reduction in performance compared to the fixed clean reward formulation, which requires human labels of IQA. In the experiments with the TRUS data, we see a trend where after an initial rise in performance with increasing holdout set rejection ratio, the performance slightly drops, this may be a dataset-specific phenomenon, since such drop was not observed in the lesion segmentation task on mpMR images. In the lesion segmentation task, after an initial rise, the performance seems to plateau with increasing holdout set rejection ratio. Nevertheless, some explanations for the plateau or small decrease may include high variance in the predictions, limited possible performance improvement due to the use of overall quality-controlled data from clinical trials, and over-fitting of the controller.

As summarised in Fig. 6a and 7, the disagreement between the learnt task-specific IQA, for the prostate lesion segmentation task, and task-agnostic IQA, for the mpMR images, shows that learning varying definitions of IQA is possible within the proposed framework. It is interesting that while the task performance for the task-agnostic IQA was not comparable with the task-specific IQA, it was still able to offer improved performance compared to a non-selective baseline. This is potentially because it may be more difficult to perform any task on images which have a large amounts of defects.

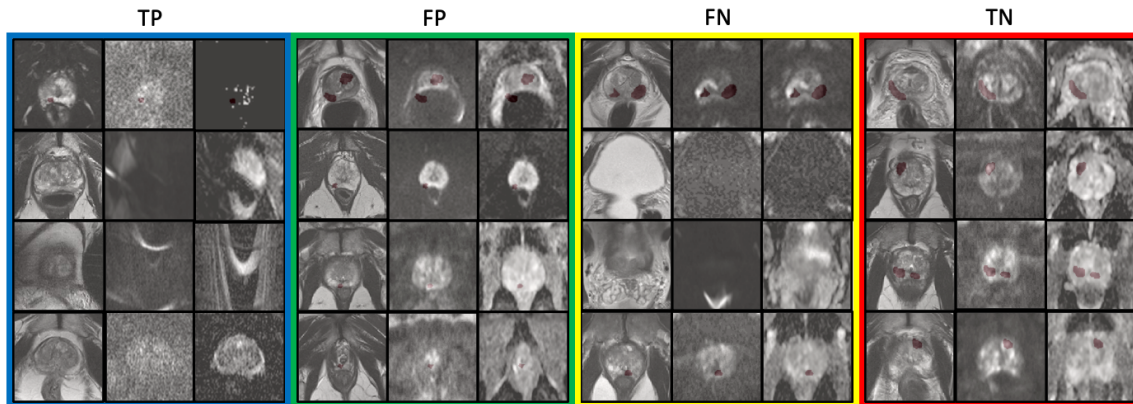


Figure 7: Samples of mpMR images from the lesion segmentation task (see Fig. 6a for definitions of TP, FN, FP and TN). In each coloured box, from left to right, the three channels of the same slice are shown (channels are T2-weighted MR, diffusion-weighted MR, and apparent diffusion coefficient MR). **TP** (Blue): Four slices form four different 3D volumes showing visible susceptibility and grain artefacts (Mazaheri et al., 2013) (valued low by both the task-specific IQA and the task-agnostic IQA); **FP** (green): Four slices form four different 3D volumes showing a misaligned sample (third slice) and some samples with a very small lesion size with low tissue density contrast between lesion and surroundings (valued low by the task-specific IQA and high by the task-agnostic IQA); **FN** (yellow): Two slices from the one 3D volume (first and second slice) and two slices from another 3D volume (third and fourth slices) where the region which contains the lesion has no apparent artefacts but distortion artefacts are present in other slices (valued high by the task-specific IQA and low by the task-agnostic IQA); **TN** (red): Samples with no apparent artefacts (valued high by both the task-specific IQA and the task-agnostic IQA). It is also interesting to note that for the shaped reward formulation with $\phi = 0.90$, the controller predicted values for these samples can be ordered from low to high in the following order: TP, FP, FN, TN (with samples at the top, in the same coloured box, being lower valued compared to samples on the bottom, if from two different 3D volumes).

Manual adjustment of the trade-off between the task-specific and task-agnostic IQA, using the proposed reward shaping strategy, allows for different IQA definitions to be learnt which may be useful under different clinical scenarios. For given a scenario, a learnt definition can then be used to obtain relevant IQA scores for new samples. As an example, to inform re-acquisition decisions, it may be useful to set a threshold, on the controller-predicted IQA scores, where the IQA controller was trained using a shaped reward, such that newly acquired images that impact task performance negatively due to artefacts may be identifiable as opposed to identifying all samples that negatively impact performance regardless of the cause (when using fully task-specific IQA). In Fig. 7, the ranking of the samples based on IQA with $\phi = 0.90$ shows that it may be possible to define such a threshold using a holdout set rejection ratio such that only samples that impact task performance due to image quality defects such as grain, susceptibility and misalignment may be flagged for re-acquisition. The shaped reward thus provides a means to identify samples with quality

defects that impact task-performance. This is in contrast to fully task-specific IQA which identifies all samples that impact task performance, regardless of the cause, and also to fully task-agnostic IQA which identifies samples that have quality defects, regardless of impact on task performance.

The classification of samples presented in Fig. 7 shows that the task-specific IQA and task-agnostic IQA learn to value samples differently. The low-quality samples flagged by the fully task-specific IQA, i.e. TP and FP, appear to either have imaging defects including distortion, grain or susceptibility, or appear to be clinically challenging e.g. samples with small lesions. It should be noted that, however, on its own the fully task-specific IQA cannot distinguish between samples with artefacts and those that are clinically challenging. The low-quality samples flagged by the fully task-agnostic IQA, i.e. TP and FN, appear to either contain artefacts within (TP) or outside (FN) regions of interest such that they do not appear to impact the target task. Contrastingly, when we observe the valuation based on $\phi = 0.90$ we see that, in FP, samples with quality defects also impacting the target task (TP), e.g. problems within regions of interest, are valued lowest; and sequence-misaligned samples are valued lower compared to clinically challenging samples. These are examples indicating that setting a threshold on the controller with shaped reward is effective, such that samples with defects impacting the target task can be identified for re-acquisition.

While the overlapping measures of IQA do not achieve the best average performance, they provide a means to identify samples for which performance may be improved by re-acquisition or defect correction, such as artefact removal or de-noising. This is because re-acquisition of samples with defects that do not impact the task is not beneficial and potentially expensive (e.g. samples with artefacts outside regions of interest). On the other hand, re-acquisition of clinically challenging samples is futile since performance for these samples cannot be improved by re-acquisition (e.g. samples with small tumour size). By definition, the fully task-specific IQA achieves highest performance (assessing based on simple sample removal), since it removes both clinically challenging samples and samples which impact the task due to imaging defects, without any distinction between the two. In contrast, the overlapping measure with the shaped reward can identify samples that impact the target task due to potentially correctable imaging defects such that if these defects were to be corrected, a performance improvement may be seen for the particular samples. For re-acquisition decisions, the overlapping measure also provides a possibility to identify samples such that target task performance, for those samples, may be improved if they are re-acquired.

It is also interesting that the task-agnostic IQA was able to identify samples which have quality defects such as distortion but which do not impact the target task due to being out of plane of the tumour. While this information may not be directly useful for the lesion segmentation task, it provides insight into types of imaging parameters or protocols, that are less important to a specific diagnosis task, for a more efficient and streamlined clinical implementation in the future.

It is important to highlight that the overall ranking of task amenability on a set of images will be altered using the combinatory IQA considering both types of qualities, compared with that from a purely task-specific IQA. The potential alternative strategy would be using the task-agnostic IQA on the subset of images selected by the task-specific IQA, with respect to a pre-defined threshold on task-specific image quality, or vice versa for a different

potential application. Adjusting the reward shaping hyper-parameter at training time, using the scheme proposed in this work, is capable of achieving equivalent selections without the need for per-sample adjustment of thresholds. Adjusting these thresholds may also be inefficient during training time when individual or small batch of images are assessed. The combinatory IQA can, therefore, be used to output controller scores in a single forward pass, for new samples, where a threshold can be specified at training to produce re-acquisition decisions. Using either measure separately or in two sequential stages requires adjustment of holdout set rejection ratio thresholds for both the task-agnostic and task-specific qualities, which in addition to potentially requiring per-sample adjustment of the thresholds, would also require two forward passes through two separate controllers.

In this work, in addition to summarising the framework to learn task-specific IQA, previously presented in Saeed et al. (2021), we have presented a mechanism which allows for task-agnostic IQA to be learnt without any human labels of quality. Moreover, the reward shaping mechanism is proposed with a manually adjustable trade-off between the task-specific and task-agnostic IQA which may be tuned for a wide range of potential applications. These extended methodologies were evaluated using a diagnostic target task of prostate lesion segmentation using mpMR images acquired from clinical prostate cancer patients.

Acknowledgments

This work was supported by the International Alliance for Cancer Early Detection, an alliance between Cancer Research UK [C28070/A30912; C73666/A31378], Canary Center at Stanford University, the University of Cambridge, OHSU Knight Cancer Institute, University College London and the University of Manchester. This work is supported by the Wellcome/EPSRC Centre for Interventional and Surgical Sciences [203145Z/16/Z]. F. Giganti is a recipient of the 2020 Young Investigator Award funded by the Prostate Cancer Foundation / CRIS Cancer Foundation. Z.M.C. Baum is supported by the Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarships-Doctoral Program, and the University College London Overseas and Graduate Research Scholarships. Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R37CA260346. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Ethical Standards

The work follows appropriate ethical standards in conducting research and writing the manuscript, following all applicable laws and regulations regarding treatment of animals or human subjects.

Conflicts of Interest

We declare that we do not have conflicts of interest.

References

- A. H. Abdi, C. Luong, T. Tsang, G. Allan, S. Nouranian, J. Jue, D. Hawley, S. Fleming, K. Gin, J. Swift, R. Rohling, and P. Abolmaesumi. Automatic quality assessment of echocardiograms using convolutional neural networks: Feasibility on the apical four-chamber view. *IEEE Transactions on Medical Imaging*, 36(6):1221–1230, 2017. doi: 10.1109/TMI.2017.2690836.
- M. Ahdoot, A. H. Lebastchi, B. Turkbey, B. Wood, and P. A. Pinto. Contemporary treatments in prostate cancer focal therapy. *Current opinion in oncology*, 31(3):200, 2019.
- H. Ahmed, A. Freeman, A. Kirkham, M. Sahu, R. Scott, C. Allen, J. Van der Meulen, and M. Emberton. Focal therapy for localized prostate cancer: a phase i/ii trial. *The Journal of urology*, 185(4):1246–1255, 2011.
- H. U. Ahmed, A. E.-S. Bosaily, L. C. Brown, R. Gabe, R. Kaplan, M. K. Parmar, Y. Collaco-Moraes, K. Ward, R. G. Hindley, A. Freeman, et al. Diagnostic accuracy of multi-parametric mri and trus biopsy in prostate cancer (promis): a paired validating confirmatory study. *The Lancet*, 389(10071):815–822, 2017.
- Z. Baum, E. Bonmati, L. Cristoni, A. Walden, F. Prados, B. Kanber, D. Barratt, D. Hawkes, G. Parker, C. Wheeler-Kingshott, and Y. Hu. Image quality assessment for closed-loop computer-assisted lung ultrasound. In C. A. Linte and J. H. Siewerdsen, editors, *Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 11598, pages 160 – 166. International Society for Optics and Photonics, SPIE, 2021. doi: 10.1117/12.2581865.
- A. E.-S. Bosaily, C. Parker, L. Brown, R. Gabe, R. Hindley, R. Kaplan, M. Emberton, H. Ahmed, P. Group, et al. Promis—prostate mr imaging study: a paired validating cohort study evaluating the role of multi-parametric mri in men with clinical suspicion of prostate cancer. *Contemporary clinical trials*, 42:26–40, 2015.
- R. Cao, X. Zhong, S. Shakeri, A. M. Bajgirani, S. A. Mirak, D. Enzmann, S. S. Raman, and K. Sung. Prostate cancer detection and segmentation in multi-parametric mri via cnn and conditional random field. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1900–1904, 2019. doi: 10.1109/ISBI.2019.8759584.
- J. Chen, Z. Wan, J. Zhang, W. Li, Y. Chen, Y. Li, and Y. Duan. Medical image segmentation and reconstruction of prostate tumor based on 3d alexnet. *Computer Methods and Programs in Biomedicine*, 200:105878, 2021. ISSN 0169-2607. doi: <https://doi.org/10.1016/j.cmpb.2020.105878>. URL <https://www.sciencedirect.com/science/article/pii/S0169260720317119>.
- M. Y. Chen, M. A. Woodruff, P. Dasgupta, and N. J. Rukin. Variability in accuracy of prostate cancer segmentation among radiologists, urologists, and scientists. *Cancer Medicine*, 9(19):7172–7182, 2020. doi: <https://doi.org/10.1002/cam4.3386>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cam4.3386>.

- M. K. Choong, R. Logeswaran, and M. Bister. Improving diagnostic quality of mr images through controlled lossy compression using spiht. *Journal of Medical Systems*, 30(3): 139–143, 2006.
- L. Chow and R. Paramesran. Review of medical image quality assessment. *Biomed. Signal Processing and Control*, 27:145 – 154, 2016.
- L. Chow and H. Rajagopal. Comparison of difference mean opinion score (dmos) of magnetic resonance images with full-reference image quality assessment (fr-iqa). In *Image Processing, Image Analysis and Real-Time Imaging (IPARTI) Symposium*, 2015.
- E. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. Le. Autoaugment: Learning augmentation policies from data, 2019.
- S. J. Daly. Visible differences predictor: an algorithm for the assessment of image fidelity. In *Human Vision, Visual Processing, and Digital Display III*, volume 1666, pages 2–15. International Society for Optics and Photonics, 1992.
- H. Davis, S. Russell, E. Barriga, M. Abramoff, and P. Soliz. Vision-based, real-time retinal image quality assessment. In *2009 22nd IEEE Int. Symp. on Computer-Based Medical Systems*, pages 1–6, 2009.
- A. De Angelis, A. Moschitta, F. Russo, and P. Carbone. Image quality assessment: an overview and some metrological considerations. In *2007 IEEE International Workshop on Advanced Methods for Uncertainty Estimation in Measurement*, pages 47–52. IEEE, 2007.
- M. de Rooij, B. Israël, T. Barrett, F. Giganti, A. Padhani, V. Panebianco, J. Richenberg, G. Salomon, I. Schoots, G. Villeirs, et al. Focus on the quality of prostate multiparametric magnetic resonance imaging: synopsis of the esur/esui recommendations on quality assessment and interpretation of images and radiologists’ training. *European Urology: Official Journal of the European Association of Urology*, 2020.
- L. Dickinson, H. U. Ahmed, A. Kirkham, C. Allen, A. Freeman, J. Barber, R. G. Hindley, T. Leslie, C. Ogden, R. Persad, et al. A multi-centre prospective development study evaluating focal therapy using high intensity focused ultrasound for localised prostate cancer: the index study. *Contemporary clinical trials*, 36(1):68–80, 2013.
- O. Dietrich, J. G. Raya, S. B. Reeder, M. F. Reiser, and S. O. Schoenberg. Measurement of signal-to-noise ratios in mr images: influence of multichannel coils, parallel imaging, and reconstruction filters. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 26(2):375–385, 2007.
- J. Dutta, S. Ahn, and Q. Li. Quantitative statistical methods for image quality assessment. *Theranostics*, 3(10):741, 2013.
- B. L. Eck, R. Fahmi, K. M. Brown, S. Zabic, N. Raihani, J. Miao, and D. L. Wilson. Computational and human observer image quality evaluation of low dose, knowledge-based ct iterative reconstruction. *Medical physics*, 42(10):6098–6111, 2015.

- S. Esses, X. Lu, T. Zhao, K. Shanbhogue, B. Dane, M. Bruno, and H. Chandarana. Automated image quality evaluation of t2-weighted liver mri utilizing deep learning architecture. *Journal of Magnetic Resonance Imaging*, 47(3):723–728, 2018.
- M. Fuderer. The information content of mr images. *IEEE Transactions on Medical Imaging*, 7(4):368–380, 1988.
- A. Geissler, A. Gartus, T. Foki, A. R. Tahamtan, R. Beisteiner, and M. Barth. Contrast-to-noise ratio (cnr) as a quality parameter in fmri. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 25(6):1263–1270, 2007.
- N. Ghavami, Y. Hu, E. Gibson, E. Bonmati, M. Emberton, C. M. Moore, and D. C. Barratt. Automatic segmentation of prostate mri using convolutional neural networks: Investigating the impact of network architecture on the accuracy of volume measurement and mri-ultrasound registration. *Medical Image Analysis*, 58:101558, 2019. ISSN 1361-8415.
- F. Giganti, C. Allen, M. Emberton, C. M. Moore, V. Kasivisvanathan, P. S. Group, et al. Prostate imaging quality (pi-qual): a new quality control scoring system for multiparametric magnetic resonance imaging of the prostate from the precision trial. *European urology oncology*, 3(5):615–619, 2020.
- L. Gondara. Medical image denoising using convolutional denoising autoencoders. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pages 241–246, 2016. doi: 10.1109/ICDMW.2016.0041.
- S. Hamid, I. A. Donaldson, Y. Hu, R. Rodell, B. Villarini, E. Bonmati, P. Tranter, S. Punwani, H. S. Sidhu, S. Willis, et al. The smarttarget biopsy trial: a prospective, within-person randomised, blinded trial comparing the accuracy of visual-registration and magnetic resonance imaging/ultrasound image-fusion targeted biopsies for prostate cancer risk stratification. *European urology*, 75(5):733–740, 2019.
- M. C. Hemmsen, M. M. Petersen, S. I. Nikolov, M. B. Nielsen, and J. A. Jensen. Ultrasound image quality assessment: A framework for evaluation of clinical image quality. In *Medical Imaging 2010: Ultrasonic Imaging, Tomography, and Therapy*, volume 7629, page 76290C. International Society for Optics and Photonics, 2010.
- R. M. Henkelman. Measurement of signal intensities in the presence of noise in mr images. *Medical physics*, 12(2):232–233, 1985.
- D. Huo, D. Xu, Z.-P. Liang, and D. Wilson. Application of perceptual difference model on regularization techniques of parallel mr imaging. *Magnetic Resonance Imaging*, 24(2):123–132, 2006.
- Y. Jiang, D. Huo, and D. L. Wilson. Methods for quantitative image quality evaluation of mri parallel reconstructions: detection and perceptual difference model. *Magnetic Resonance Imaging*, 25(5):712–721, 2007.

- M. M. Kalayeh, T. Marin, and J. G. Brankov. Generalization evaluation of machine learning numerical observers for image quality assessment. *IEEE transactions on nuclear science*, 60(3):1609–1618, 2013.
- L. Kaufman, D. M. Kramer, L. E. Crooks, and D. A. Ortendahl. Measuring signal-to-noise ratios in mr imaging. *Radiology*, 173(1):265–267, 1989.
- I. Kowalik-Urbaniak, D. Brunet, J. Wang, D. Koff, N. Smolarski-Koff, E. R. Vrscay, B. Wallace, and Z. Wang. The quest for ‘diagnostically lossless’ medical image compression: a comparative study of objective quality metrics for compressed medical images. 9037:903717, 2014.
- A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. *NeurIPS*, 2012.
- B. Kumar, G. Sinha, and K. Thakur. Quality assessment of compressed mr medical images using general regression neural network. *International Journal of Pure & Applied Sciences & Technology*, 7(2), 2011.
- R. Kumar and M. Rattan. Analysis of various quality metrics for medical image processing. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2(11):137–144, 2012.
- S. N. Lavasani, A. Mostaar, and M. Ashtiyani. Automatic prostate cancer segmentation using kinetic analysis in dynamic contrast-enhanced mri. *Journal of biomedical physics & engineering*, 8(1):107, 2018.
- Y. Li, H. Zhang, J. Chen, P. Song, J. Ren, Q. Zhang, and K. Jia. Non-reference image quality assessment based on deep clustering. *Signal Processing: Image Communication*, 83:115781, 2020. ISSN 0923-5965. doi: <https://doi.org/10.1016/j.image.2020.115781>.
- Z. Liao, H. Girgis, A. Abdi, H. Vaseli, J. Hetherington, R. Rohling, K. Gin, T. Tsang, and P. Abolmaesumi. On modelling label uncertainty in deep neural networks: automatic estimation of intra-observer variability in 2d echocardiography quality assessment. *IEEE Transactions on Medical Imaging*, 39(6):1868–1883, 2019.
- T. Lillicrap, J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra. Continuous control with deep reinforcement learning, 2019.
- Z. Lin, S. Li, D. Ni, Y. Liao, H. Wen, J. Du, S. Chen, T. Wang, and B. Lei. Multi-task learning for quality assessment of fetal head ultrasound images. *Medical Image Analysis*, 58:101548, 2019. ISSN 1361-8415. doi: <https://doi.org/10.1016/j.media.2019.101548>.
- M. Linch, G. Goh, C. Hiley, Y. Shanmugabavan, N. McGranahan, A. Rowan, Y. Wong, H. King, A. Furness, A. Freeman, et al. Intratumoural evolutionary landscape of high-risk prostate cancer: the progeny study of genomic and immune parameters. *Annals of Oncology*, 28(10):2472–2480, 2017.

- C. P. Loizou, C. S. Pattichis, M. Pantziaris, T. Tyllis, and A. Nicolaides. Quality evaluation of ultrasound imaging in the carotid artery based on normalization and speckle reduction filtering. *Medical and Biological Engineering and Computing*, 44(5):414–426, 2006.
- S. Marshall and S. Taneja. Focal therapy for prostate cancer: The current status. *Prostate International*, 3(2):35–41, 2015. ISSN 2287-8882. doi: <https://doi.org/10.1016/j.prn.2015.03.007>. URL <https://www.sciencedirect.com/science/article/pii/S228788821500015X>.
- Y. Mazaheri, H. Vargas, G. Nyman, O. Akin, and H. Hricak. Image artifacts on prostate diffusion-weighted magnetic resonance imaging: trade-offs at 1.5 tesla and 3.0 tesla. *Academic radiology*, 20 8:1041–7, 2013.
- S. W. Merriel, G. Funston, and W. Hamilton. Prostate cancer in primary care. *Advances in therapy*, 35(9):1285–1294, 2018.
- J. Miao, D. Huo, and D. L. Wilson. Quantitative image quality evaluation of mr images using perceptual difference models. *Medical physics*, 35(6Part1):2541–2553, 2008.
- B. Mortamet, M. A. Bernstein, C. R. Jack Jr, J. L. Gunter, C. Ward, P. J. Britson, R. Meuli, J.-P. Thiran, and G. Krueger. Automatic quality assessment in structural brain magnetic resonance imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 62(2):365–372, 2009.
- I. Oksuz, J. R. Clough, B. Ruijsink, E. P. Anton, A. Bustin, G. Cruz, C. Prieto, A. P. King, and J. A. Schnabel. Deep learning-based detection and correction of cardiac mr motion artefacts during reconstruction for high-quality segmentation. *IEEE Transactions on Medical Imaging*, 39:4001–4010, 2020.
- C. Orczyk, D. Barratt, C. Brew-Graves, Y. Peng Hu, A. Freeman, N. McCartan, I. Potyka, N. Ramachandran, R. Rodell, N. R. Williams, et al. Prostate radiofrequency focal ablation (proraf) trial: A prospective development study evaluating a bipolar radiofrequency device to treat prostate cancer. *The Journal of Urology*, 205(4):1090–1099, 2021.
- D. Racine, A. H. Ba, J. G. Ott, F. O. Bochud, and F. R. Verdun. Objective assessment of low contrast detectability in computed tomography with channelized hotelling observer. *Physica Medica*, 32(1):76–83, 2016.
- D. K. S. Rangaraju, K. Kumar, and C. Renumadhavi. Review paper on quantitative image quality assessment—medical ultrasound images. *Int. J. Eng. Research and Technology*, 1(4), 2012.
- A. Razumov, O. Y. Rogov, and D. V. Dylov. Optimal mri undersampling patterns for ultimate benefit of medical vision tasks. *arXiv preprint arXiv:2108.04914*, 2021.
- O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, volume 9351. Springer, 2015.

- O. Rouvière, P. Puech, R. Renard-Penna, M. Claudon, C. Roy, F. Mège-Lechevallier, M. Decaussin-Petrucci, M. Dubreuil-Chambardel, L. Magaud, L. Remontet, et al. Use of prostate systematic and targeted biopsy on the basis of multiparametric mri in biopsy-naive patients (mri-first): a prospective, multicentre, paired diagnostic study. *The Lancet Oncology*, 20(1):100–109, 2019.
- S. U. Saeed, Y. Fu, Z. M. C. Baum, Q. Yang, M. Rusu, R. E. Fan, G. A. Sonn, D. C. Barratt, and Y. Hu. Learning image quality assessment by reinforcing task amenable data selection. In A. Feragen, S. Sommer, J. Schnabel, and M. Nielsen, editors, *Information Processing in Medical Imaging*, pages 755–766, Cham, 2021. Springer International Publishing.
- K. A. Salem, J. S. Lewin, A. J. Aschoff, J. L. Duerk, and D. L. Wilson. Validation of a human vision model for image quality evaluation of fast interventional magnetic resonance imaging. *Journal of Electronic Imaging*, 11(2):224–235, 2002.
- J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. Proximal policy optimization algorithms, 2017.
- Y.-H. Shiao, T.-J. Chen, K.-S. Chuang, C.-H. Lin, and C.-C. Chuang. Quality of compressed medical images. *Journal of Digital Imaging*, 20(2):149, 2007.
- Y. Shima, A. Suwa, Y. Gomi, H. Nogawa, H. Nagata, and H. Tanaka. Qualitative and quantitative assessment of video transmitted by dvts (digital video transport system) in surgical telemedicine. *Journal of telemedicine and telecare*, 13(3):148–153, 2007.
- L. A. Simmons, A. Kanthabalan, M. Arya, T. Briggs, D. Barratt, S. C. Charman, A. Freeman, D. Hawkes, Y. Hu, C. Jameson, et al. Accuracy of transperineal targeted prostate biopsies, visual estimation and image fusion in men needing repeat biopsy in the picture trial. *The Journal of urology*, 200(6):1227–1234, 2018.
- A. Sinha, P. Malo, and K. Deb. A review on bilevel optimization: From classical to evolutionary approaches and applications. *IEEE Transactions on Evolutionary Computation*, 22:276–295, 2018.
- S. J. C. Soerensen, R. E. Fan, A. Seetharaman, L. Chen, W. Shao, I. Bhattacharya, Y.-H. Kim, R. Sood, M. Borre, B. I. Chung, K. J. To’o, M. Rusu, and G. A. Sonn. Deep learning improves speed and accuracy of prostate gland segmentations on magnetic resonance imaging for targeted biopsy. *The Journal of urology*, 206(3):604–612, September 2021. ISSN 0022-5347. doi: 10.1097/ju.0000000000001783.
- P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning, ICML ’08*, page 1096–1103, New York, NY, USA, 2008. Association for Computing Machinery. ISBN 9781605582054. doi: 10.1145/1390156.1390294. URL <https://doi.org/10.1145/1390156.1390294>.
- P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, 11(12), 2010.

- J. Wang, J. Miao, X. Yang, R. Li, G. Zhou, Y. Huang, Z. Lin, W. Xue, X. Jia, J. Zhou, R. Huang, and D. Ni. Auto-weighting for breast cancer classification in multimodal ultrasound. In *MICCAI 2020*, pages 190–199, Cham, 2020. Springer. ISBN 978-3-030-59725-2.
- Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4): 600–612, 2004.
- J. C. Weinreb, J. O. Barentsz, P. L. Choyke, F. Cornud, M. A. Haider, K. J. Macura, D. Margolis, M. D. Schnall, F. Shtern, C. M. Tempany, et al. Pi-rads prostate imaging–reporting and data system: 2015, version 2. *European urology*, 69(1):16–40, 2016.
- J. P. Woodard and M. P. Carley-Spencer. No-reference image quality metrics for structural mri. *Neuroinformatics*, 4(3):243–262, 2006.
- L. Wu, J. Cheng, S. Li, B. Lei, T. Wang, and D. Ni. Fuiqa: Fetal ultrasound image quality assessment with deep convolutional networks. *IEEE Trans. on Cybernetics*, 47(5):1336–1349, 2017.
- R. Yang, J. Su, and W. Yu. No-reference image quality assessment based on deep learning method. In *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, pages 476–479, 2017. doi: 10.1109/ITOEC.2017.8122340.
- J. Ye, Y. Xue, L. R. Long, S. Antani, Z. Xue, K. C. Cheng, and X. Huang. Synthetic sample selection via reinforcement learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 53–63. Springer, 2020.
- J. Yoon, S. Arik, and T. Pfister. Data valuation using reinforcement learning, 2020.
- G. Zago, R. Andreão, B. Dorizzi, E. Ottoni, and T. Salles. Retinal image quality assessment using deep learning. *Computers in Biology and Medicine*, 103:64 – 70, 2018.
- X. Zhang, Q. Wang, J. Zhang, and Z. Zhong. Adversarial autoaugment, 2019.
- B. Zoph and Q. Le. Neural architecture search with reinforcement learning, 2017.
- Özgün Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation, 2016.