Bi-parametric prostate MR image synthesis using pathology and sequence-conditioned stable diffusion

Shaheer U. Saeed[1], Tom Syer[2], Wen Yan[1, 3], Qianye Yang[1], Mark Emberton[4], Shonit Punwani[2], Matthew J. Clarkson[1], Dean C. Barratt[1], Yipeng Hu[1]

[1] CMIC, WEISS, Dept. Medical Physics, University College London;
[2] CMI, University College London;
[3] Dept. Electrical Engineering, City University of Hong Kong;
[4] Dept. Urology, University College London Hospital; Division of Surgery and Interventional Science, University College London

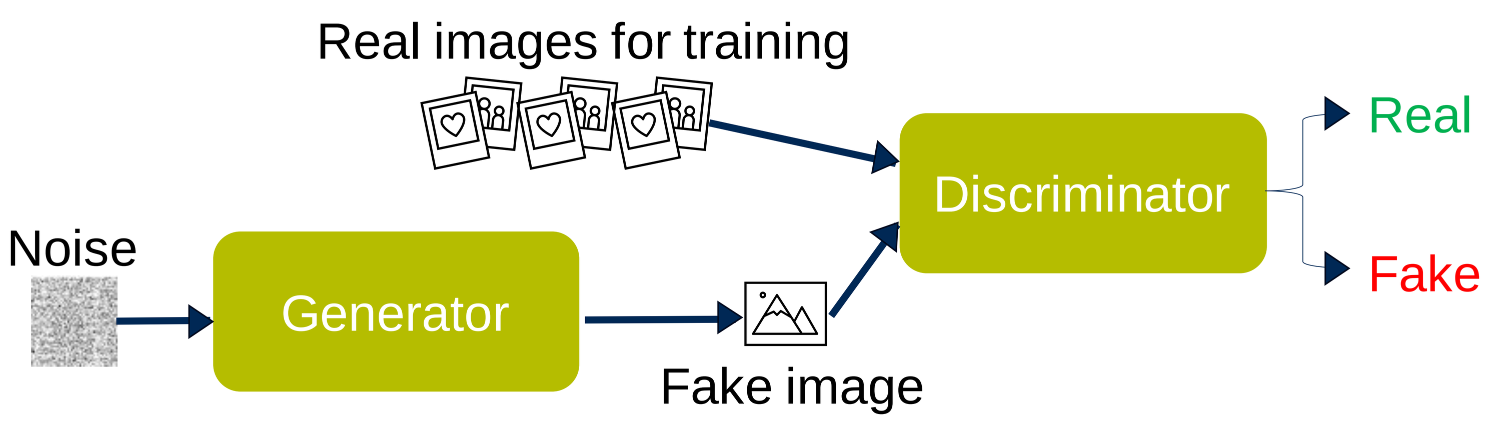
UCL

What is diffusion-based image synthesis?

- A newly proposed conditional image synthesis method that can accommodate complicated conditions e.g., long text prompts
- "Inspired by considerations from non-equilibrium thermodynamics" (Ho et al 2020) - gradually add noise to data in a Markov chain and the model learns to reverse this
- Models are trained with billions of image-text pairs form the web e.g.:

Generative Adversarial Networks (GANs) for image synthesis

- Two functions:
  Generator
  Discriminator
- Noise -> Generator -> Fake Image -> Discriminator
  Noise fed to the generator produces a fake image which is scored on its realism by the discriminator
- Realism prediction by discriminator serves as a supervision signal for training both functions
- Try to go directly from noise to the original image (intuitively difficult)
- Can be conditioned on other variables using those as input instead of, or along with noise
- Suffer from problems such as mode collapse (creating the same image every time), diminishing gradients etc.



Diffusion Models for image synthesis

Diffusion works based on sequentially added small amounts of noise rather than trying to go from full noise image to a sample

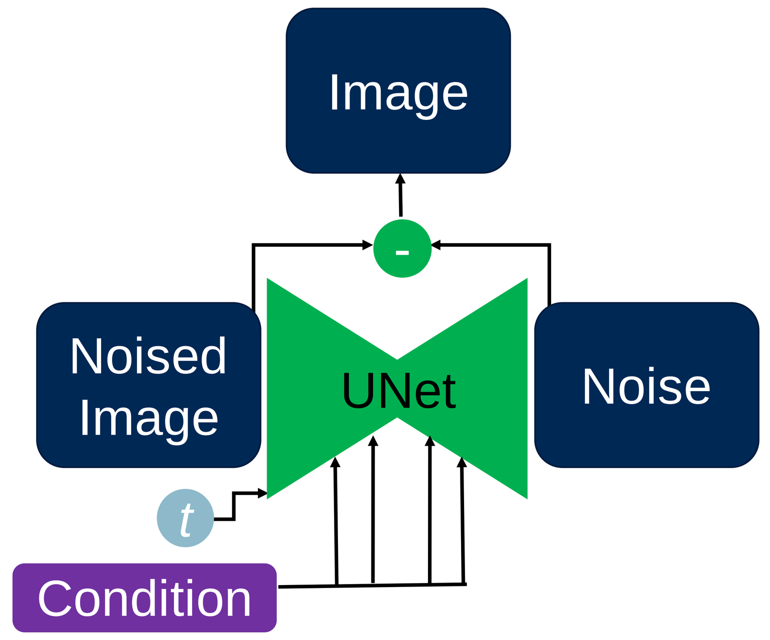
Slow to train and sample an image from due to this iterative nature

We want to learn a function that goes from the last noised image to the first one

We have paired data but this task it too difficult and results are not good (GANs attempt this)

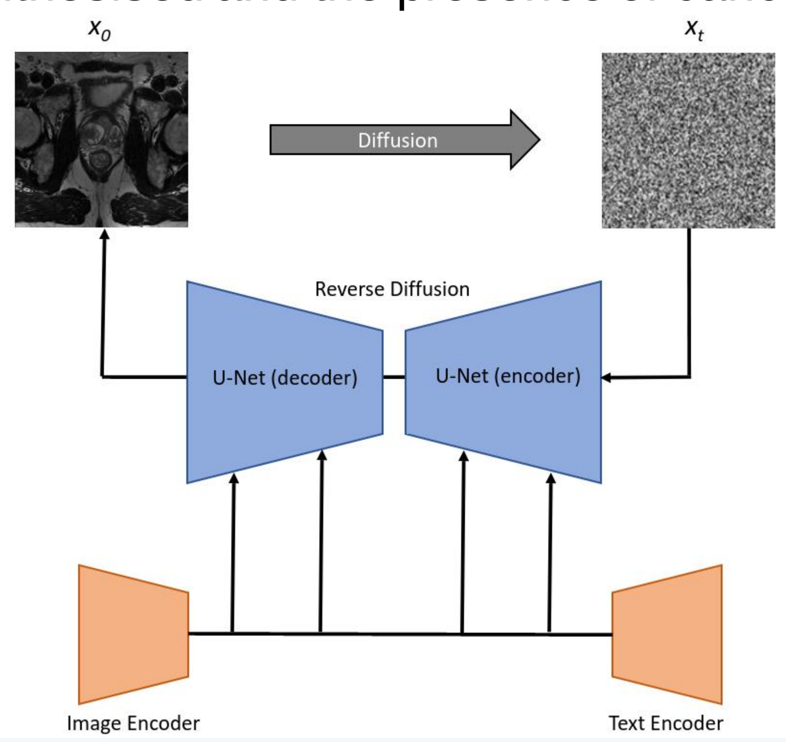
So, we only learn to remove some noise, given the time-step, t

Stable diffusion is more strongly conditioned on conditioning variables e.g., text prompts due to 'classifier free guidance':



Our proposed method

We condition the model on images as well as text to control the MR sequence synthesised and the presence of cancer, respectively



Experiments

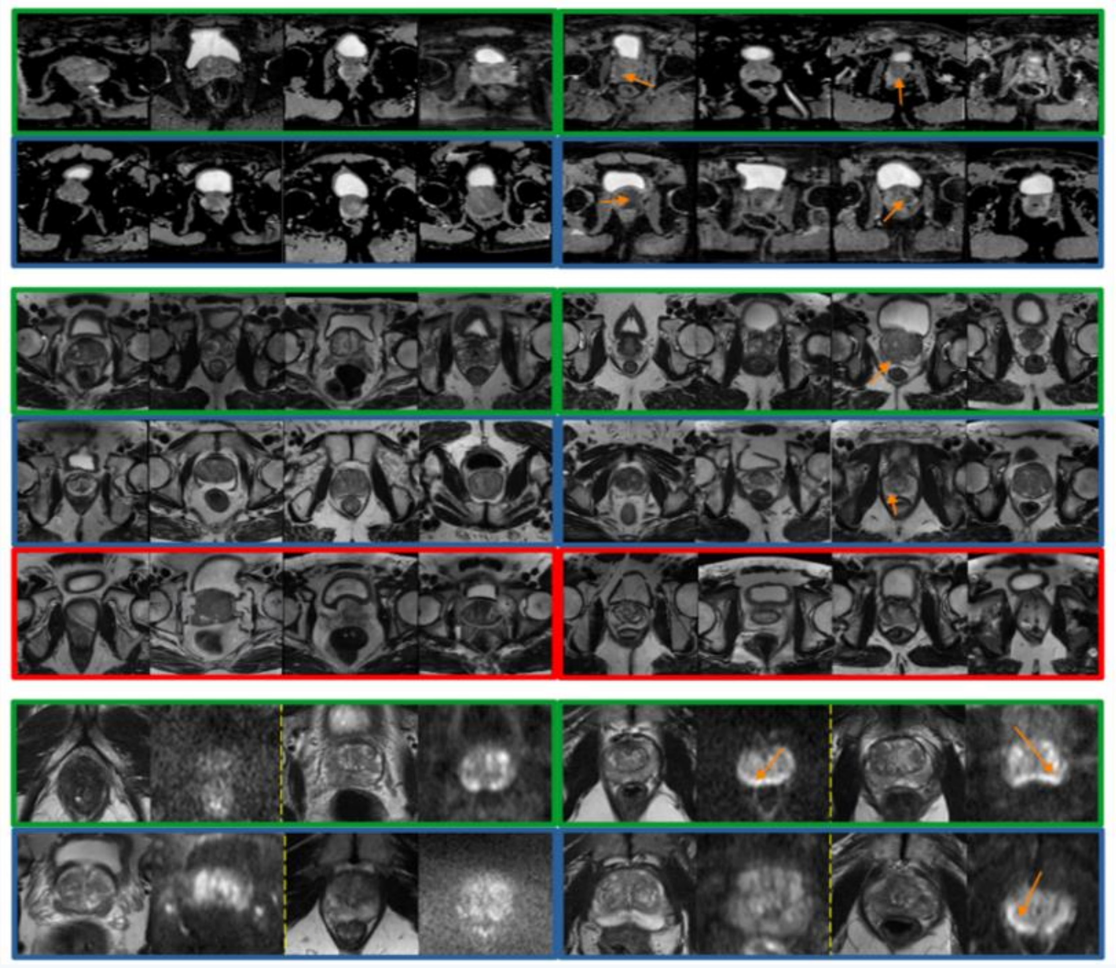
Conditioned on: MR sequence, lesion presence and images for paired bi-parametric MR synthesis

Results (Quantitative)

Expert clinician was only able to identify synthesised images from real ones with an average accuracy of 0.59

Lesion identification accuracy as performed by the expert is 0.60 for real images and 0.63 for synthesised images

ML model for lesion identification on real images improved by 5.8% accuracy, when trained with real data augmented with syntesised, compared to real only



Results (Qualitative)

Green: Real
Blue: Fake

Left: No Lesion
Right: Lesion

Top: ADC
Middle: T2W
Bottom: Paired T2W-DW

Conclusion

We present a stable-diffusion-based image synthesis method together with a conditioning scheme to generate realistic prostate MR images

The sequence of MR, presence of lesions and synthesis of paired data may be controlled

Quantitative results demonstrate the usability of the generated images for use cases such as training trainee clinicians/ radiologists and for improving machine learning model task performance.

shaheer.saeed.17@ucl.ac.uk

UKRI Engineering and Physical Sciences Research Council
wellcome EPSRC centre interventional+ surgical sciences
CANCER RESEARCH UK
ACED INTERNATIONAL ALLIANCE FOR CANCER EARLY DETECTION