Digital Radiology is now part of standard medical practice. X-ray images and CT scans are stored as digital files and sent over the Internet. Radiologist can often work from home using a standard internet connection, X-rays are sent routinely to analysis centers in India and computerized archiving of images is becoming the standard. This digitization of radiology resulted in very large cost saving for the healthcare industry.

The state of digital pathology is very different. In pathology, glass slides are still the dominating standard. Almost all analysis is performed using a physical slide, consultation with an off-site specialist requires mailing physical slides and storing of the physical slides is the dominant archival method. The fundamental reason is that, unlike radiology, high magnification of the image is needed in order to produce a reliable diagnosis. While it is now possible to use scanning microscopes to image a whole slide at high magnification, the resulting files are huge. Scanning an area of 1cm X 1cm at a 40X magnification results in a 50,000 X 50,000 pixels image, which requires 20 Giga-byte of disk space. Using lossy compression methods reduces this to 1Giga-byte, but might loose critical information. In either case, moving files of that size around the Internet is slow and expensive. The final result is that digital pathology is not widely adopted.

However, it is very instructive to study the way in which pathologists evaluate a slide. They do not scan it at high magnification. Doing so would take too much time. Instead, they study the slide at a low magnification and identify locations that are suspicious. They then pan the slide to those locations and study them at maximal magnification. Only 5-10 locations that are studied at high magnification. Storing only the images that the pathologist studies (the low magnification image and the 5-10 high magnification images) reduces the required disk space to around 10 mega-byte, a 1000-fold reduction in space compared with the original image.

Over the last two years we have been developing an artificial intelligence method for identifying the regions in a prostate biopsy slide that a pathologist would want to observe in high magnification in order to diagnose prostate cancer. Our method is a combination of computer vision algorithm for detecting important visual elements in the image, such as glands, and a machine learning algorithm. A pathologist (Dr Stephen Baird) identifies the regions in low magnification images that he finds suspicious and would like to see in high magnification. The machine learning algorithm processes this information and generates a rule that mimics the choices made by Dr Baird. Our method is reaching a good level of performance. The next step we plan to take is to present the summarized images to other pathologists and evaluate whether this restricted view of the slide degrades the accuracy of their diagnosis.