With the rapid development of nano-scale materials syntheses in the last 30 years for applications such as catalysis or energy storage technologies, there is an increasing need to develop high-performance microscopic methods to image and characterise these samples. The field of scanning or transmission electron microscopy (STEM), where the electron beams used have a wavelength 100x smaller than conventional light, has evolved in response to this need.

Using conventional photographic cameras in our daily lives we are already familiar with the way dark settings reduces the sharpness of the resulting images. Similarly, for STEM, if we want sharp images of nanostructures we need bright settings, this means having a high intensity electron beam. For many materials the beam intensity which yields high fidelity images is intense enough to severely damage the material. Using weaker electron beams to combat this brings a trade-off in data-quality. Algorithms have been developed that statistically enhance images attained when using lower intensity electron beams. However, current state-of-the-art techniques can take minutes to enhance a single image or introduce unacceptable image-border artefacts. This poses several key issues:

1. Having to wait minutes for feedback between minor adjustments in parameters such as focus, specimen position or beam intensity is prohibitively inefficient in terms of characterising the material since the sample may have moved or degraded in this time, leading to a reduction in throughput. The additional time spent using the electron microscopes is expensive (up to €2,000/day in operating costs)
2. The computational complexity of current state-of-the-art algorithms often requires supercomputers to execute, hence real-time image improvement is presently out of the question.
3. The introduction of image-border artefacts reduces the fidelity of the data, confidence in the obtained results, and makes image montaging (stitching) impossible. This impedes imaging utility especially in biological imaging (amongst others) and limits imaging uptake in these disciplines.

These issues pose the following research questions:

1. Can an algorithm capable of real-time enhancement of electron microscopy images be created, orders of magnitude times faster than what is currently available?
2. Could such an algorithm also require less computational resources (so that it can be sustainably deployed on the existing infrastructure without additional computing costs)?
3. What is an acceptable fidelity compromise if this speed-improvement can be achieved?

Successful creation of such an algorithm would enable characterisation of materials otherwise too fragile to withstand the minimum intensity electron beam needed to produce useful images. This includes novel two-dimensional (2D) materials, biological materials, materials for solar-cells and energy storage, and solid-state catalysts such as metal-organic frameworks (MOFs).

Preliminary work in this direction was performed last year as part of an unpaid research experience internship. A basic framework of a solution was proposed which functions as follows; the image is divided into smaller patches so as to stay within the capacity of fast access RAM. These patches are parsed to a matrix where principal component analysis may be performed. This “orthogonalises” the data, organising the resulting descriptors of the data (eigenvalues) in order of their descriptive ability (magnitude). These mathematical operations borrow expertise from other disciplines of computing where years of optimisation has led to near perfect implementation of these operations. These methods do not pose the present bottleneck.

There comes a point where the eigenvalues below a certain magnitude no longer contribute useful information to the image, and instead contribute only noise. If a distinction between this useful/non-useful information can be reliably and automatically determined (without slow manual human intervention), then useful information can be preserved while efficiently rejecting random noise.

The identification of this cut off point is where I will focus most of my research efforts. Once the valuable information in the data has been preserved the image is reconstructed and patched together. The reconstruction phase is still not completely optimised and requires further study. The image enhancement must be performed very quickly with the additional goal of it being computable using only the RAM of a PC; no supercomputers required.

I have two main research goals. The first and smaller of which is to optimise the re-patching phase. I have a concept for a sub-algorithm which should be capable of speeding up this phase by up to 80%. Rather than pasting reconstructed image patches back onto the image canvas and letting patches overlap in the process, my algorithm predetermines the overlap and, by making careful use of high-dimensional data structures, reconstructs the image on a strip-by-strip rather than a patch-by-patch basis, massively reducing the number of operations. I plan to fully integrate my re-patching algorithm, including the capability of dynamic determination of the suitability of either my algorithm or Prof. Jones' for re-patching of a given image. Performance will be benchmarked using simulated electron microscopy data (where the ground-truth is known, quantitative quality analysis is possible), and then later using fresh experimental data recorded using the Nion UltraSTEM200 microscope.

The second and more exploratory research goal is to investigate the potential of Machine Learning (ML) in determining the optimal number of eigenvalues to employ. Eigenvalue magnitude alone is insufficient to determine this number, therefore eigenvalues are plotted together with combinations of other calculated parameters (patch entropy, isotropy between successive patches, covariance, etc.) so that the data segregates into two distinct clusters that may make choosing the cut-off point clearer. No combination of parameters has yet been found that cause such a clear segregation. ML may help solve this problem by identifying more complicated, less intuitive, yet highly descriptive parameters across which the data varies as we would hope. My involvement will be ML implementation and ensuring the balance is struck between the ability of the resulting output labels to cluster the data distinctly, and computational efficiency of calculating a given output label.

In conclusion, my research objective is to co-create, with Prof. Lewys Jones, an image enhancement algorithm suited to low intensity electron beams fast enough for real-time characterisation of nano-materials, ultimately enabling characterisation of a broader range of nano-materials.