In the modern world, computers have improved almost every aspect of our lives. Recently, home security has become the latest target of the technology revolution. Companies like Ring [RING] and Eufy [EUFY] offer IoT devices like doorbells and cameras to allow their customers to monitor their property 24/7. On top of traditional surveillance, these companies also provide software solutions to monitor the footage recorded by their devices and interpret it. For example, a doorbell may recognise who is at the front door and allow them to enter or alert the user to the presence of a stranger if it doesn’t. However, the computational intensity of these inferences means footage must be transferred from the devices to more powerful servers.

In order to preserve privacy, video is encrypted before it is sent to the server. But when the inference algorithms are executing, the footage must be decrypted. This is an immediate privacy concern. Having the ability to decrypt the footage exposes the opportunity for employees of these companies to access constant surveillance of peoples’ homes. The possibilities for exploitation are endless. Malicious actors could use this information to monitor people’s location, appraise their belongings, or use the contents of footage for extortion, to name a few. Homomorphic Encryption may provide a solution to this.

Homomorphic Encryption (henceforth HE) is a cryptographic method of encrypting data such that mathematical operations can be performed on encrypted data, or \textit{ciphertext}, itself, rather than on the raw data, or \textit{plaintext}. For example, consider the operation $3 \times 5$. In a traditional encryption scheme, the plain values $3$ and $5$ would be multiplied before encrypting the result. Using a homomorphic scheme, the $3$ and $5$ can be encrypted, and the ciphertexts multiplied so that when the ciphertext is decrypted, the plaintext is $15$. But can this technique be scaled to more complex algorithms, like those required for surveillance?

More specifically, is it possible to extract the moving objects from a frame of HE video data? Moving object detection, also known as \textit{foreground extraction} or \textit{background subtraction}, is fundamental to modern surveillance systems. Detecting when, for example, somebody enters a property, allows the security systems to alert their owners, possibly pre-empting a break-in. To perform this analysis, the contents of a video must be modelled using a, usually probabilistic, function that allows significant changes in a pixels’ value to be discerned. The difficulty of this arises when accounting for environmental changes which will cause numerical variation, such as light levels when moving from day to night or different weather conditions causing objects to distort.