Computers have improved almost every aspect of modern life. 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 analyse footage. For example, a doorbell may recognise a visitor or alert to the presence of a stranger. However, the computational intensity of inference means footage must be transferred to more powerful servers.

To preserve security, video is encrypted before transmission to the server. However, the footage must be decrypted when the inference algorithms are executing. 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. Consequently, malicious actors could use this information to monitor peoples’ location, appraise their belongings, use the contents of footage for extortion, and more. Homomorphic Encryption (henceforth HE) may provide a solution to this.

In cryptography, HE describes encryption schemes that allow mathematical operations to be performed directly on encrypted data, or \textit{ciphertext}, rather than on raw data, or \textit{plaintext}. For example, consider $3 \times 5$. 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 $15$. An open question is, 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 is fundamental to surveillance. Detecting when, for example, somebody enters a property allows security systems to alert their owners, possibly pre-empting a break-in. However, more than just motion must be sensed. Objects must be extracted and analysed to prevent users being notified of unimportant events like, for example, leaves blowing onto a property. Gradual changes and random noise in an environment make modelling a background for object detection a significant challenge to overcome.