2019

Cross-correlation Template Matching for Liver Localisation in Computed Tomography

Patrick Leyden
University College Dublin

Martin O'Connell
University College Dublin

Derek Greene
University College Dublin

See next page for additional authors

Recommended Citation

Leydon, P., O'Connell, M., Green, D., & Curran, K. (2019). Cross-correlation template matching for liver localisation in computed tomography. *IMVIP 2019: Irish Machine Vision & Image Processing*, Technological University Dublin, Dublin, Ireland, August 28-30. doi:10.21427/8fgf-y086

Cross-Correlation Template Matching for Liver Localisation in Computed Tomography

Mr Patrick Leydon, Dr Martin O'Connell, Dr Derek Greene, and Dr Kathleen Curran

University College Dublin, Dublin 4, Ireland

Abstract

Many of the current approaches to automatic organ localisation in medical imaging require a large amount of labelled patient data to *train* systems to accurately identify specific anatomical features. Cross-Correlation, also known as template matching, is a statistical method of assessing the similarity between a template image and a target image. This method has been modified and presented here to localize the liver in Computed Tomography volume images in the Coronal and Sagital planes to achieve a mean positioning error of approximately 11 mm and 20 mm respectively based on between 1 and 25 datasets to create the template liver.

Keywords: Medical Imaging, Template Matching, Cross-Correlation, Localisation, Liver.

1 Introduction

Organ localisation has many applications in medical image analysis, segmentation, multimodal image registration as well as improving outcomes for interventions that rely on image guidance [Venkatraghavan and Ranjan, 2011]. Advances in Machine Learning (ML), and Deep Learning (DL) in particular, have dominated the literature in recent years [Shen et al., 2017]. Often in order to successfully *train* a deep network for an imaging task a very large number of examples are needed. This proved to be a challenge for DL applications in medical imaging.

The introduction of Convolutional Neural Networks (CNNs) as a method of DL have demonstrated an ability to achieve good results in classification tasks on relatively small datasets typical of medically based studies and there are many examples of the successful application of CNNs in organ identification and localisation.

The method builds several layers of low level features from a limited amount of training images (or 3D volumes) and uses these features to classify organs, lesions, etc via a learned function which is optimized during training [Xu et al., 2019]. Despite the lower number of datasets needed for CNNs in comparison with more traditional DL approaches, the difference being *~100's* as opposed to *~1000's*, the data requirement may often still be a limitation that is difficult to overcome without access to a large hospital database. There is also a considerable time factor involved in labelling data as well the time required for successful *training* and optimization of these systems.

In such instances more traditional image processing techniques may be the best option as they tend to require substantially less data than ML methods. Cross-Correlation is one such method that has been successfully used in image processing for many years and has been used here to localize the liver from a series of CT datasets.

1.1 Cross-Correlation

Cross-Correlation, also referred to as *template matching*, is a statistically based method of assessing the similarity between a *reference* image matrix and a *target* image matrix. The reference is typically a subset, or subregion, of the target. The similarity is assessed at every location as the reference is moved across the target.

Cross-Correlation refers specifically to how the similarity is measured. [Sarvaiya et al., 2009]. As the reference matrix steps across the target matrix the element-by-element product is calculated and summed. This summed value is the similarity measure at that particular location of the target.

Once Cross-Correlation has been performed for each pixel of the target the maximum value, or *brightest pixel*, corresponds to the location where the reference best matches the target. Often prior to performing the cross-correlation it is necessary to pad the target image to get results at image edges and to normalize or appropriately adjust the range of values used in both target and reference to obtain accurate results.

Other methods, such as Sum of Squared Differences, also exist but these methods have been shown to be less robust than Cross-Correlation to variance in the target image [Mahmood and Khan, 2011].

2 Methods

The CT data is a 3D volume of the patient built up from a series of 2D *slices* along the Axial direction. The pixel, or *voxel*, intensities in a CT relate to the physical density of tissues with high values corresponding to high density tissues, such as bone, and lower values relating to low density tissues, such as those in the lungs. The 3D volume can be projected in the Coronal and Sagital planes to give 2D images of the original volume.

The method outlined here creates a 2D template image of the liver based on the averaging of several different images of the liver from several different patients. This is the template used for Cross-Correlation performed on separate patient target datasets (those *not* used to create the template) and location of the maximum pixel value corresponds to the most likely location of the liver in the target images. By locating the liver in the Coronal and Sagital planes its complete 3D position can be established.

For analysis the dataset was randomized prior to creating the template liver image from a range of images (n = 1 - 25). At each stage the process was repeated 10 times to reduce the impact of including potentially biased images in either templates liver images or target CT projections. This was performed, and analysed, separately for both the Sagital and Coronal projections.

2.1 Data

After receiving ethical approval by the institution, PET-CT scans of 55 patients was compiled. The data was anonymised prior to analysis to ensure compliance with data protection procedures. The scans took place between October 2012 and November 2017 and were all performed on the same Siemens Biograph 16 PETCT scanner. The mean patient age at the time of the scan was 55.7 (*26 - 85*) years with a mean weight of 79.4 (*44 - 127*) kgs. All scans used in the study were oncology related, with all patient referrals indicating a history of cancer. The scan reports were reviewed to ensure the patient liver was considered to be healthy and had not received any recent therapy or intervention which could alter the appearance of the images.

A large portion of the scans were for melanoma (14) and lymphoma (6) with the remainder of scans being for conditions such as, myeloma, breast, testicular, head & neck and lung cancer.

2.2 Preprocessing

All data processing and analysis was carried out using Matlab software. CT images are usually displayed in Hounsfield Units (HU) which can be achieved by applying the calibration information available in the metadata of the dicom files used as standard in medical imaging. This has the effect of centring pixel intensity values around 0 HU which corresponds to the density of water, and are typically within the range of between -1000 HU to +3000 HU for air and cortical bone, respectively [Bushberg and Boone, 2011]. The data was upsampled from (1 x 1 x 3) mm^3 to (1 x 1 x 1) mm^3 to improve visualisation along the z-axis and to make distance calculations more straight forward.

In order to emphasise the appearance of the liver in the images a pixel based thresholding was applied to retain pixel values that relate to liver and other soft tissues only, this was between 0 HU and 150 HU

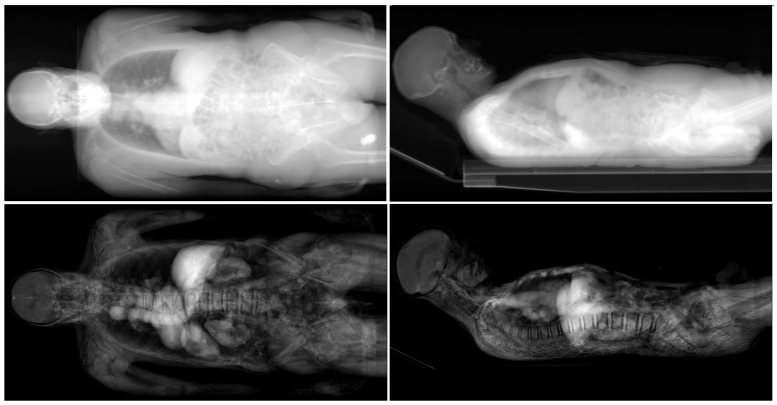

Figure 1: Shows an example of Coronal and Sagital projections before and after thresholding between 0 and 150 HU, demonstrating enhanced appearance of liver and other soft tissues.

[Lamba et al., 2014]. After this the projections were performed. Projections were pixel summations along the desired axis. The results of this thresholding and projection can be seen in Figure 1. The projection of the coronal image was from posterior to anterior and hence the liver appears on the patient left.

2.2.1 Labelling

Once all 55 datasets were thresholded and projected the resulting 110 images were manually labelled using a *bounding box* to enclose the liver. The bounding box sizes and locations were stored for later use in creating liver templates and as the *Ground Truth* for assessment of localisation accuracy.

2.3 Liver Templates

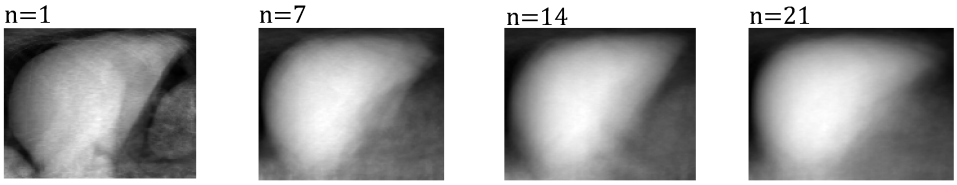

Figure 2: Shows examples of the coronal liver template images with the number of images used for each indicated.

Liver templates which were used for Cross-Correlation were a composite of between 1 and 25 individual images of the liver created during the labelling of the projection images. In order to get a useful template image the dimensions of each image had to match, and so the images were resized to the overall mean size of all the images used to form the template. Examples of some of the resulting templates can be seen in Figure 2.

2.4 Template Matching

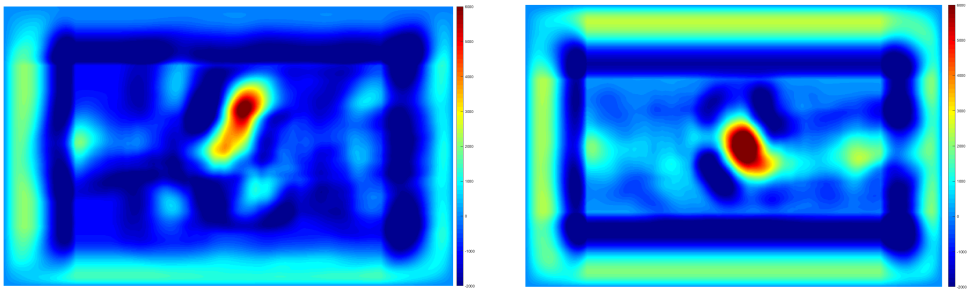

Figure 3: Shows the result of running Cross-Correlation on Coronal and Sagital projections.

Before running the Cross-Correlation both the template liver and target projection images were adjusted to be in the range of -1 and +1. This was done in order to maximize similarity output by not only counting areas which matched but also penalizing areas that did not match.

Cross-Correlation, described previously, of the liver templates was used to determine the most likely location of liver the in the target image. The location of maximum pixel value was used as the centre point of the bounding box the dimensions of which correspond to those of the template used. Figure 3 demonstrates the result of running a cross-correlation between a template liver and target CT projections. The square differences between the locations of maximum pixel value and the centre point of the bounding box established during the earlier labelling session were used to assess positioning accuracy.

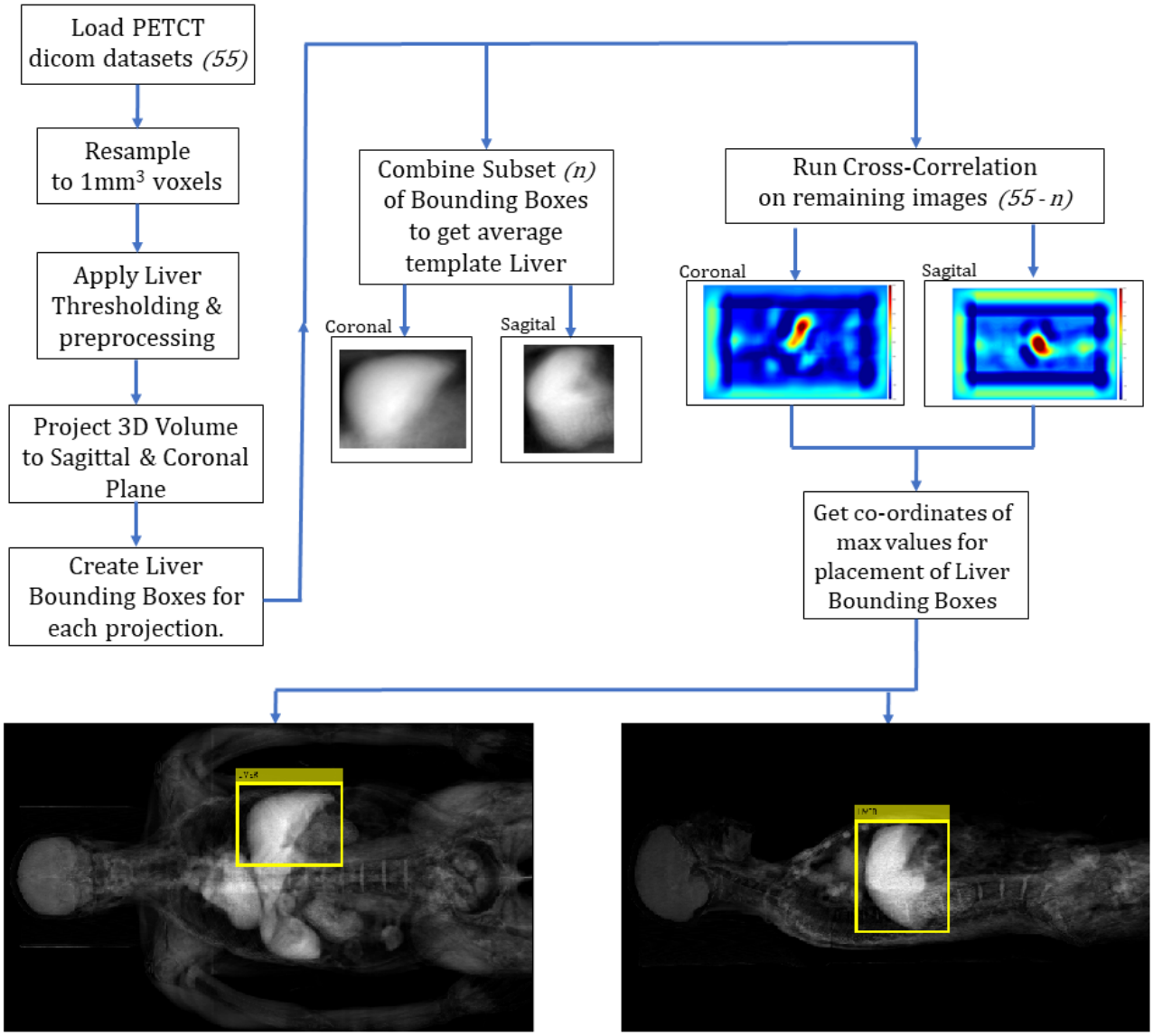

Figure 4: Shows the steps taken for the liver localisation procedure.

3 Results

The results are presented in the graphs shown in Figure 5. For all 10 randomizations the overall mean of the Square Error between the centre point of the ground truth bound box and the bounding box positioned using Cross-Correlation was ~140.1 and ~481.0 for the Coronal and Sagital planes respectively. Each pixel

represented a distance of 1 mm and so these errors correspond to distances of approximately 11 mm and 20 mm.

The inclusion of more images (n = 1 - 25) when creating the liver template reduced the overall error significantly up to approximately 10 individual images. With the largest improvement seen when any more than 1 image was used to create the template. The error seems to plateau after approximately 10 with only small fluctuations evident thereafter. The error in the Coronal images appears to begin to increase from n = 20 on.

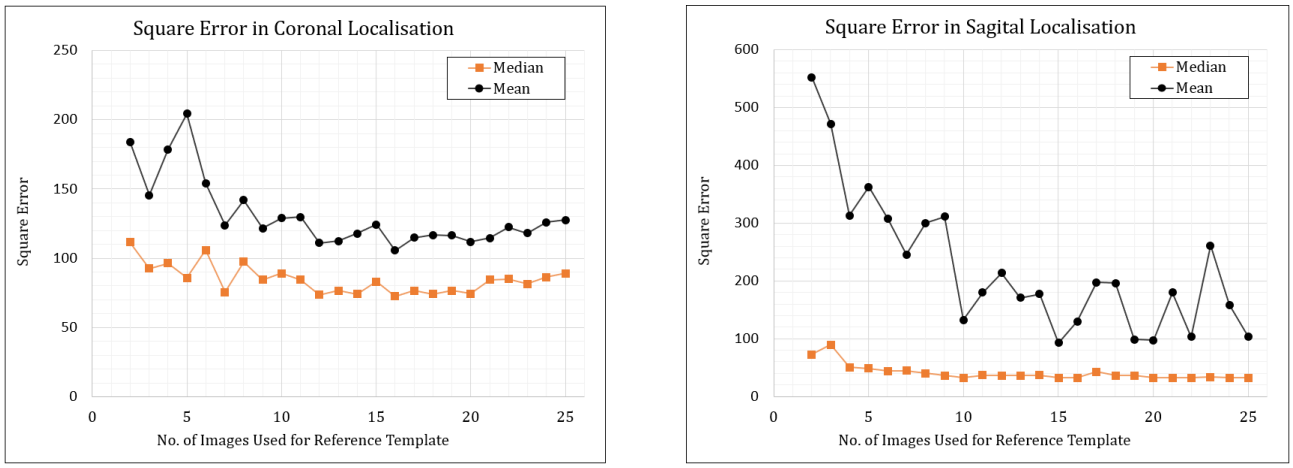

Figure 5: The results of Square Errors between Cross-Correlation localisation method and Ground Truth.

4 Conclusions

The Cross-Correlation method presented performed well in the liver localisation task. The majority of positioning errors were less than 1 cm which is comparable to other non-DL methods such as regression forests [Criminisi et al., 2013]. Images that demonstrated higher errors tended to be for patients with unusually large, or unusually shaped livers. The reliance of the templates on an *average* liver meant that positioning errors will be inevitable for a certain percentage of scans. However the inclusion of an ensemble of liver templates of varying sizes could help reduce this error in the future. The fact that this process required only a very limited number of datasets (~10s) means that the technique may be useful in situations where the amount of data available is too small or the time requirement is too limiting to attempt a DL solution.

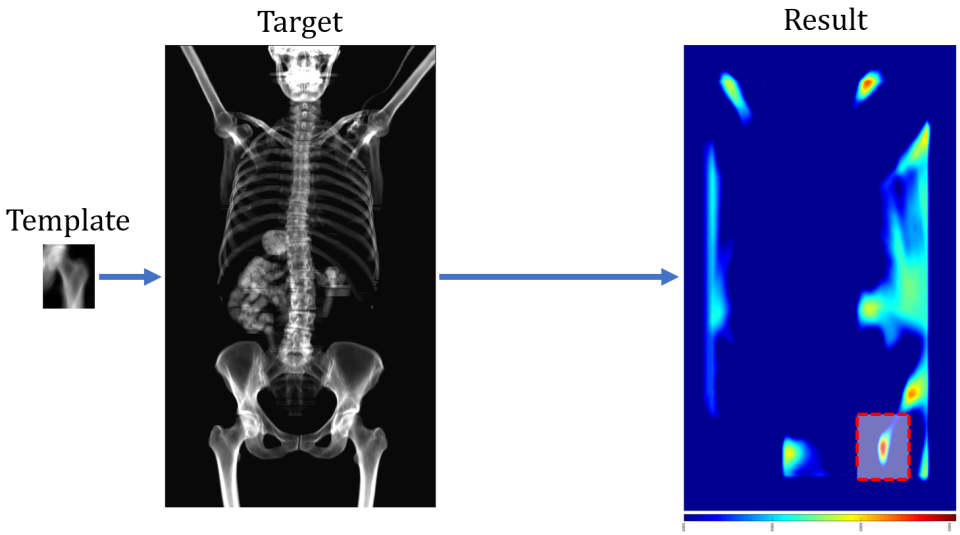

Figure 6: Shows application of the technique to the cortical bone of the proximal femur.

Although the method presented was used for localisation of the liver it could also be successfully applied to

other anatomical regions and tissues where there is a need to identify organs quickly and accurately. Following a similar process to that outlined above demonstrated similar results for localisation of the proximal femur from a CT dataset as shown in Figure 6.

The PET component of the scans was not utilised in as part of this study but it could be incorporated to automatically return objective quantitative physiological data on glucose metabolism within specific organs and tissues in the body and reduce the subjectivity associated with the standard clinical measurement technique [Büyükdereli et al., 2016].

References

[Bushberg and Boone, 2011] Bushberg, J. T. and Boone, J. M. (2011). *The essential physics of medical imaging*. Lippincott Williams & Wilkins.

[Büyükdereli et al., 2016] Büyükdereli, G., Güler, M., and Şeydaoğlu, G. (2016). Interobserver and intraobserver variability among measurements of fdg pet/ct parameters in pulmonary tumors. *Balkan medical journal*, 33(3):308–315.

[Criminisi et al., 2013] Criminisi, A., Robertson, D., Konukoglu, E., Shotton, J., Pathak, S., White, S., and Siddiqui, K. (2013). Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical image analysis*, 17(8):1293–1303.

[Lamba et al., 2014] Lamba, R., McGahan, J. P., Corwin, M. T., Li, C.-S., Tran, T., Seibert, J. A., and Boone, J. M. (2014). Ct hounsfield numbers of soft tissues on unenhanced abdominal ct scans: variability between two different manufacturers' mdct scanners. *American Journal of Roentgenology*, 203(5):1013–1020.

[Mahmood and Khan, 2011] Mahmood, A. and Khan, S. (2011). Correlation-coefficient-based fast template matching through partial elimination. *IEEE Transactions on image processing*, 21(4):2099–2108.

[Sarvaiya et al., 2009] Sarvaiya, J. N., Patnaik, S., and Bombaywala, S. (2009). Image registration by template matching using normalized cross-correlation. In *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pages 819–822. IEEE.

[Shen et al., 2017] Shen, D., Wu, G., and Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19:221–248.

[Venkatraghavan and Ranjan, 2011] Venkatraghavan, V. and Ranjan, S. (2011). Generic framework for organ localization in ct and mr images. In *2011 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–7. IEEE.

[Xu et al., 2019] Xu, X., Zhou, F., Liu, B., Fu, D., and Bai, X. (2019). Efficient multiple organ localization in ct image using 3d region proposal network. *IEEE transactions on medical imaging*.