



Classifying coke using CT scans and landmark multidimensional scaling

Keith Nesbitt¹ · Fayeem Aziz¹ · Merrick Mahoney¹ · Stephan Chalup¹ · Bishnu P. Lamichhane¹

Received: 2 December 2021 / Revised: 24 November 2022 / Accepted: 16 January 2023
© The Author(s) 2023

Abstract

One factor that limits development of fundamental research on the influence of coke microstructure on its strength is the difficulty in quantifying the way that microstructure is both classified and distributed in three dimensions. To support such fundamental studies, this study evaluated a novel volumetric approach for classifying small (approx. $450 \mu\text{m}^3$) blocks of coke microstructure from 3D computed tomography scans. An automated process for classifying microstructure blocks was described. It is based on Landmark Multi-Dimensional Scaling and uses the Bhattacharyya metric and k-means clustering. The approach was evaluated using 27 coke samples across a range of coke with different properties and reliably identified 6 ordered class of coke microstructure based on the distribution of voxel intensities associated with structural density. The lower class (1–2) subblocks tend to be dominated by pores and thin walls. Typically, there is an increase in wall thickness and reduced pore sizes in the higher classes. Inert features are also likely to be seen in higher classes (5–6). In general, this approach provides an efficient automated means for identifying the 3D spatial distribution of microstructure in CT scans of coke.

Keywords Coke · Microstructure · Clustering · Classification · Computer tomography

1 Introduction

Metallurgical coke plays a key role in the blast furnace ironmaking process – it must maintain the permeability of the furnace so liquid iron may be drained from the furnace and combustion air injected into the base of the furnace (Riley 2007; Tiwari et al. 2015). To maintain permeability the coke must maintain a good size distribution and shape while being subjected to mechanical loads, high temperatures, attack by oxidising gases and dissolution in the blast furnace liquids. Important properties of the coke therefore include the strength of the coke to support the ferrous burden and prevent collapse of the reactive zones to maintain permeability in the blast furnace, as well as the reactive behaviour of coke. Optimising the blast furnace ironmaking process requires balancing the cost of making the coke with its performance in the blast furnace. Understanding how the microstructure of the coke affects its properties will significantly aid in this optimisation process.

Coke is primarily carbon and is formed when some coals are heated in the absence of air. Some of the organic molecules in coal fuse, release gas forming bubbles and on further heating resolidify to bind the non-fusible inerts into a complex 3D structure. The final structure depends on both the coal blend used in coking and the coking conditions (particle size, bulk density, and heating rate). The prediction of coke properties is difficult because of the complexity of coke microstructure, involving, pores, pore walls (from fused material), inerts and minerals, all of which may be distributed in different ways throughout the coke (Ghosh et al. 2018). Current prediction models based on empirical data often fail to work across the broad spectrum of coals found in different locations (North et al. 2018). Often heterogeneous, different 3D features can be found in different parts of the coke body (Zhang and Tao 2017). Fundamental understanding of the physical coke properties such as strength and chemical reactivity ultimately rely on better assessment of the way these 3D structures are distributed. The first step to studying the distribution of these features is an efficient way to automatically detect and classify the location and type of microstructure features in a 3D sample.

The use of 2D image analysis techniques have been used in the past to help analyse coke microstructure features

✉ Bishnu P. Lamichhane
bpl360@newcastle.edu.au

¹ School of Information and Physical Sciences, University of Newcastle, Callaghan, NSW, Australia

such as porosity, pore length, perimeter, breadth, roundness, thickness, and size distribution (Ghosh et al. 2018). These previous image-based approaches have typically studied optical 2D textures, imaged using reflected light microscopes with polished coke segments at various scales (Ghosh et al. 2018). While valuable, these approaches may miss critical 3D relationships that can help define important microstructure features related to coke quality. In terms of coke strength, these features of interest might include strong networks of pore wall connection, distributions of pore wall thickness, symmetric and non-symmetric pore properties, the presence of micro-fissures, large sized micro-textures and graphitic lamella, low numbers of contact points between inert and reactive maceral derived components (IMDC/RMDC) and regions of frothy, bubbly structures found within thin pore walls.

The broader range of factors related to coke quality and the formation of 3D carbon structure of coke have recently been reviewed, highlighting the need for more fundamental understanding of coke formation (Chen et al. 2020). Microstructure analysis on CT images is also useful in examining coke fracture using von Mises stress (Lomas et al. 2017). Our approach here is a first step towards developing analysis techniques without the need to calculate von Mises stress. There is a recent study in microstructure analysis and mineral phase transformation in blast furnace (Wu et al. 2022). We also point out a recent review paper focussing on the current characterisation methods of the coke structure (Zheng et al. 2021). Earlier, Raman spectroscopy is also applied to predict the coke quality (Rantitsch et al. 2014).

While 3D classification of microstructure may be desirable, the first problem in moving to 3D classification approaches is the practical problem of obtaining sufficient 3D volumes of coke microstructure, scanned at various scales across a range of cokes. In this research we use 27 previously scanned samples that cover a range of 9 cokes with known properties. In general, 3D volume classification has been found to be quite challenging and to require extensive training of classification models (Qi et al. 2016). Despite this, it has been successfully demonstrated for classifying small segments of structure made from sequential image segments (Zou et al. 2017; Kamnitsas et al. 2017) and automatic microstructures recognition using image segmentation (Chen et al. 2014a, b). Thus, classifiers can be trained to classify small segments or blocks of structure based on inner characteristics. Where the classes can be well defined before training 3D convolutional networks are found to be applicable (Qi et al. 2016; Zou et al. 2017). However, coke volumes typically have an irregular distribution of complex 3D structures. This can make it difficult to provide specific class samples for use in supervised classification and suggests an unsupervised classification approach might be more appropriate for microstructure features. Another approach in

3D classification is to use micro-finite element modelling, which is an effective approach to understand the relationship between the microstructure and strength (Tsafnat et al. 2008).

In this work we develop a tool that may assist in understanding the role of distributed microstructure components to overall coke strength (Fig. 1). A novel volumetric approach for classifying small (approx. $450 \mu\text{m}^3$) blocks of coke microstructure from 3D computed tomography (CT) scans is described and evaluated. An automated process for classifying microstructure blocks in turn supports the study of their distribution in a 3D sample. The final intention for this approach is to try and develop an alternative method, that characterizes cokes directly using the distribution of different classes of microstructure found in CT images of coke. It is expected that such a tool could assist in understanding the link between the distribution of microstructure types, coke strength, reactivity, and the formation of the microstructure in the plastic layer.

In this research we have developed an unsupervised classification algorithm for 3D structural features that is based on the k-means clustering (Hartigan and Wong 1979) and Landmark Multi-Dimensional Scaling (LMDS, De Silva and Tenenbaum 2004), which are applied across a range of nine different coke samples that have previously been scanned. The derived classes are further examined by using first and second order image statistics. One benefit of this approach is that it readily lends itself to the study of microstructure distribution in three dimensions (Fig. 1).

Landmark-based dimensionality reduction has been shown to improved k-means clustering performance and classification accuracy (Magdalinos et al. 2011). LMDS is a non-linear, dimensionality reduction algorithm, which is itself based on isometric feature mapping (ISOMAP) (Tenenbaum et al. 2000). This approximates the classical Multidimensional Scaling (MDS) approach (Shepard 1980; Cox and Cox 1994; Kruskal and Wish 1978). Moreover, it is computationally more efficient for large datasets. Due to the improved clustering performance and computational efficiency with the large volumetric datasets used in this work, LMDS was selected to generate a lower dimensional space for clustering regions of microstructure (De Silva and Tenenbaum 2004). This leads to a novel automatic identification of different types of 3D coke microstructure using dimension reduction and clustering techniques. To the best of our knowledge, this is done for the first time. The aim is to support further study of the distribution of microstructures at this scale and how this distribution may be related to coke strength characteristics.

To classify blocks of coke we need to compare the similarity of blocks. Therefore, as with other techniques that rely on similarity measures, an important consideration when using LMDS is the choice of distance metric (Shirkhorshidi

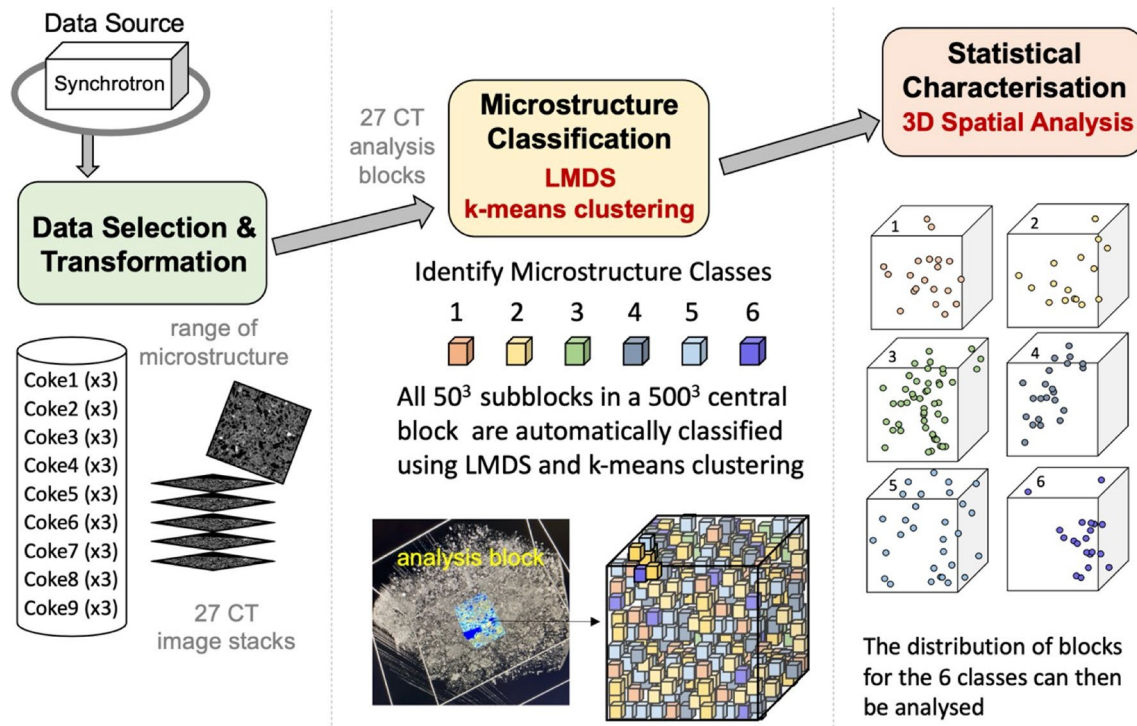


Fig. 1 An overview of the suggested process for characterising cokes based on the distribution of automatically classified blocks of microstructure. This study describes the classification step

et al. 2015; Pandit and Gupta 2011). Typically, the Euclidean metric is selected as a default metric. However, we would not expect this to be an isotropic measure, so changing the orientation of a block of coke might change the measurement outcome. To simplify this study, we would like blocks of microstructure to be classified the same regardless of the way they are orientated. We initially tested four alternative distance metrics (Chebyshev, Euclidean, Manhattan, Bhattacharyya), but of these, only the Euclidean and Bhattacharyya measures produced ordered classification results. The Bhattacharyya (Goudail et al. 2004; Chalup et al. 2007) distance metric was found to be the most reliable under rotation of coke volumes (isomorphic) and was therefore adopted as the metric for this study.

2 Samples and classification methods

2.1 Samples

The 3D volumes classified in this work were obtained from existing CT scans and include three samples from nine different cokes. These nine cokes were chosen as they cover a range of coal petrographic compositions and coke microstructure (Fig. 2, Tables 1, 2). The 27 sets of CT images were collected at the Imaging and Medical Beamline (IMBL) at

the Australian Synchrotron. The data is stored as a stack of images slices, each slice being a 2D 8-bit TIFF image. This data is acquired in a regular pattern with a regular number of pixels in each image slice and regular spacing of image slices. This generates a regular grid of voxels, each of which represents a volume of $8.893 \mu\text{m}^3$. Each voxel contains a single scalar intensity value in the range of 0 (black) to 255 (white). This intensity measure reflects the level of absorption, with darker areas in an image corresponding to pores and brighter areas to pore walls, maceral components and minerals.

While the original coke samples were irregular in shape, the volume that was analysed was constrained to a 500^3 block of voxels. This was automatically taken from the centre of the sample to ensure no irregular sample edges were included in the analysis (Fig. 3). In this study, our classification approach analyses 1000 subblocks, each subblock has 50^3 voxels. Since each voxel is approximately $9 \mu\text{m}$ in each dimension, each classified subblock of microstructure occupies a volume of approximately $450^3 \mu\text{m}^3$.

2.2 Classification methods

The first step in the classification process is Landmark Multi-Dimensional Scaling (LMDS). This process is designed to produce a 2D spatial map. With a suitable

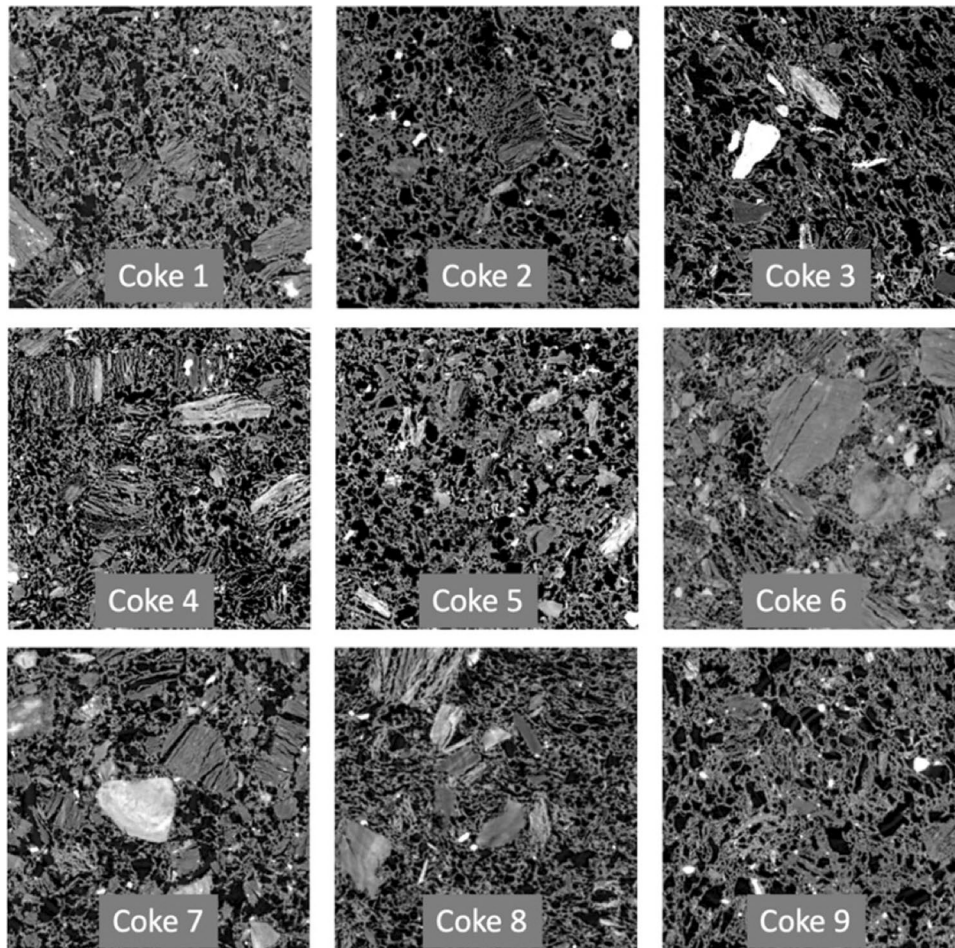


Fig. 2 Example 2D slices showing microstructure from the nine different cokes used in the study. Three samples of each coke were used

Table 1 Traditional Measures of coke quality properties (where available)

Coke No.	ASTM stability (%)	ASTM hardness (%)	Irsid I20 (%)	Irsid I10 (%)	CRI (%)	CSR (%)
1	65.6	68.1	80.6	18.2	22	71
2	66	68	81	18	22	71
3	57.3	60.6	77.4	21.3	13	72
4	65	68	78	20	24	67
5	67	71	80	18	24	69
6	62.3	65.0	77.0	22.3	40.6	51
7	n/a	n/a	n/a	n/a	n/a	n/a
8	56.9	61.0	74.9	23.4	18	69
9	22.5	69.1	56.8	21.3	n/a	n/a

distance metric, LMDS is typically used to allow 2D visualization of high-dimensional data objects. The intention is to reduce the number of dimensions that define each object so they can be plotted and compared on a two-dimensional spatial map. This is exactly the intention in this project,

to organize the blocks of coke in a 2D space so that similar blocks are close together and dissimilar blocks are far apart. Blocks that are close together can then be grouped into the same class.

The next step is to calculate the block classes using a clustering algorithm. In this project a traditional k-means Nearest Neighbours (k-NN) clustering algorithm is used. The value of k was set to six and this therefore generates six classes of microstructure. As the name suggests, the

blocks that are near neighbours in the 2D space defined by the LMDS process are allocated to the same class.

To further examine the classes, we calculated second order statistics on the volumes. We reduced each image slice in a subblock volume to 8 grey-levels and calculated

Table 2 Typical microstructure properties and coal ranks for the analyzed cokes

Coke No.	Type of coal	Microstructure properties
1	Prime coking coal	Good network of connections, good pore wall thickness, good pore properties
2	Prime coking coal	Good network of connections, good pore wall thickness, good pore properties
3	Good coking coal	Good network of connections with good pore and wall sizes
4	High rank coal	Large sized micro-textures / graphitic lamella
5	High vitrinite coal	Regions of frothy, bubbly structures with thin pore walls
6	High inerts coal	Low numbers of contact between IMDC and RMDC
7	High inerts coal	Low numbers of contact between IMDC and RMDC
8	High vitrinite coal	Regions of frothy, bubbly structures with thin pore walls
9	Low rank (high shrinkage) coal	High amount of microfissures

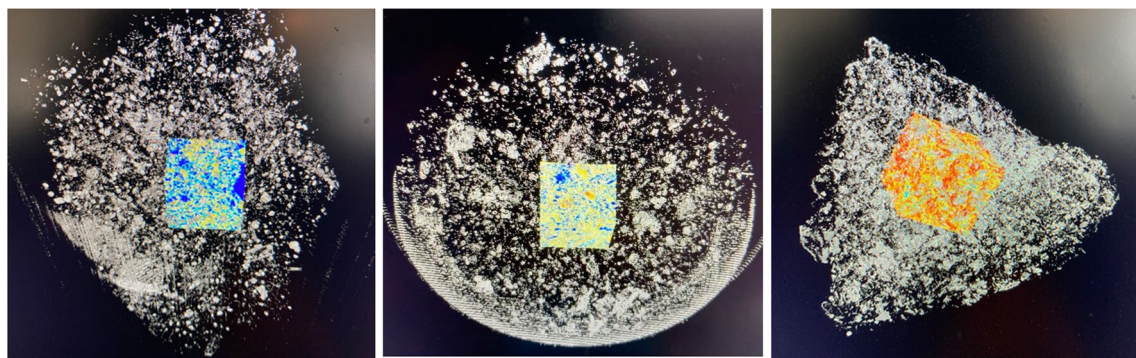


Fig. 3 Examples of how the 500³ voxel analysis blocks were automatically selected chosen from the centre of each sample

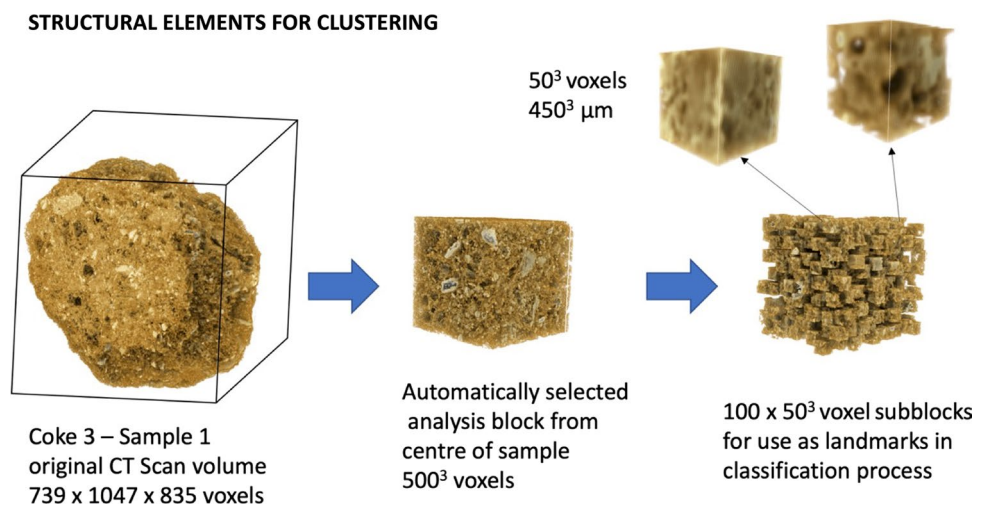


Fig. 4 Example of how subblocks of microstructure are generated for classification

a grey-level co-occurrence matrix for all pixels separated by distance of one (Materka and Strzelecki 1998). The spatial relationship of neighbouring pixels was calculated at the four angles of 0, 45, 90 and 135 degrees and results combined. However, as this only considered image slices in two dimensions (XY) we also reoriented the subblock and repeated the process to consider pixel relationships in the third dimension (XZ).

In summary, the overall approach (Figs. 3, 4) taken in this study was to

- (1) Extract a central block (500^3 voxels) from each coke sample, avoiding sample edges.
- (2) Randomly select 100 subblocks (50^3 voxels) from analysis block to act as "landmarks".
- (3) Use the randomly selected subblocks as landmarks for LMDS to generate a lower dimensional space for clustering. This step uses the Bhattacharyya distance metric (Fig. 5).
- (4) Perform k-means clustering analysis on all 1000 subblocks (50^3 voxels) in the analysis block using the lower dimensional space generated by the LMDS to determine six different classes (clusters) of microstructure (Fig. 6).
- (5) Use interactive 3D visualizations for expert review of the six generated classes (Fig. 7).
- (6) Perform additional image analysis to review features of the six classes.

3 Results

All 27 coke samples, three for each of the nine coke, were classified using LMDS and k-means clustering. Six classes of microstructure were generated for the one thousand 50^3 voxel subblocks in a standard way, by analysing a 500^3 voxel block, automatically taken from the centre of the CT scans for each sample.

The average number of classified blocks per class and coke are shown in Table 3. There was significant variation found in the way classes were distributed across samples. This variation was also evident in samples of the same coke. The heterogeneous nature of coke is well described and suggests that more samples are required to provide significant statistical evidence of variations that occur in different coke.

We analysed 3 samples of each coke, but 30 or more samples would provide better evidence. The size of our analysis block was 500^3 voxels. A larger block size may also help provide more data points. We were restricted in this study because some of our samples were too small to provide larger blocks.

In general, most of the classified subblocks (50^3 voxels) were found to belong to class 3 (43%), and the fewest blocks were found in class 1 (3%) and class 6 (1%) (Table 5). Both class 2 (21%) and class 4 (24%) contained similar numbers of subblocks, while class 5 made up 8% of the classified subblocks.

The general features of blocks found in each class were examined by experts in the domain using Drishti. Drishti is an interactive 3D visualization tool (Limaye 2012) that

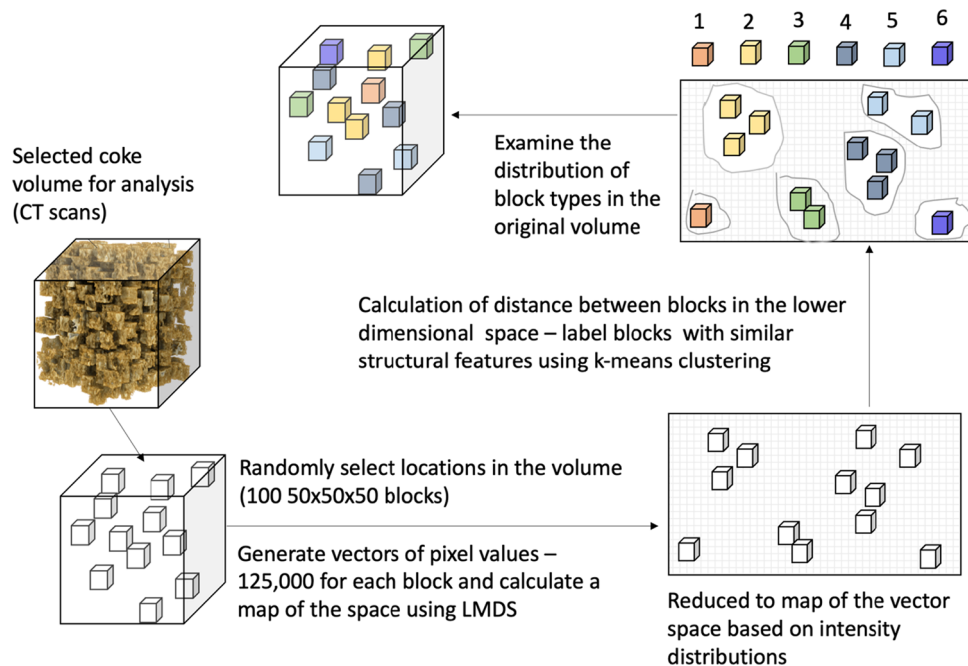


Fig. 5 An overview of the unsupervised classification approach of coke structure

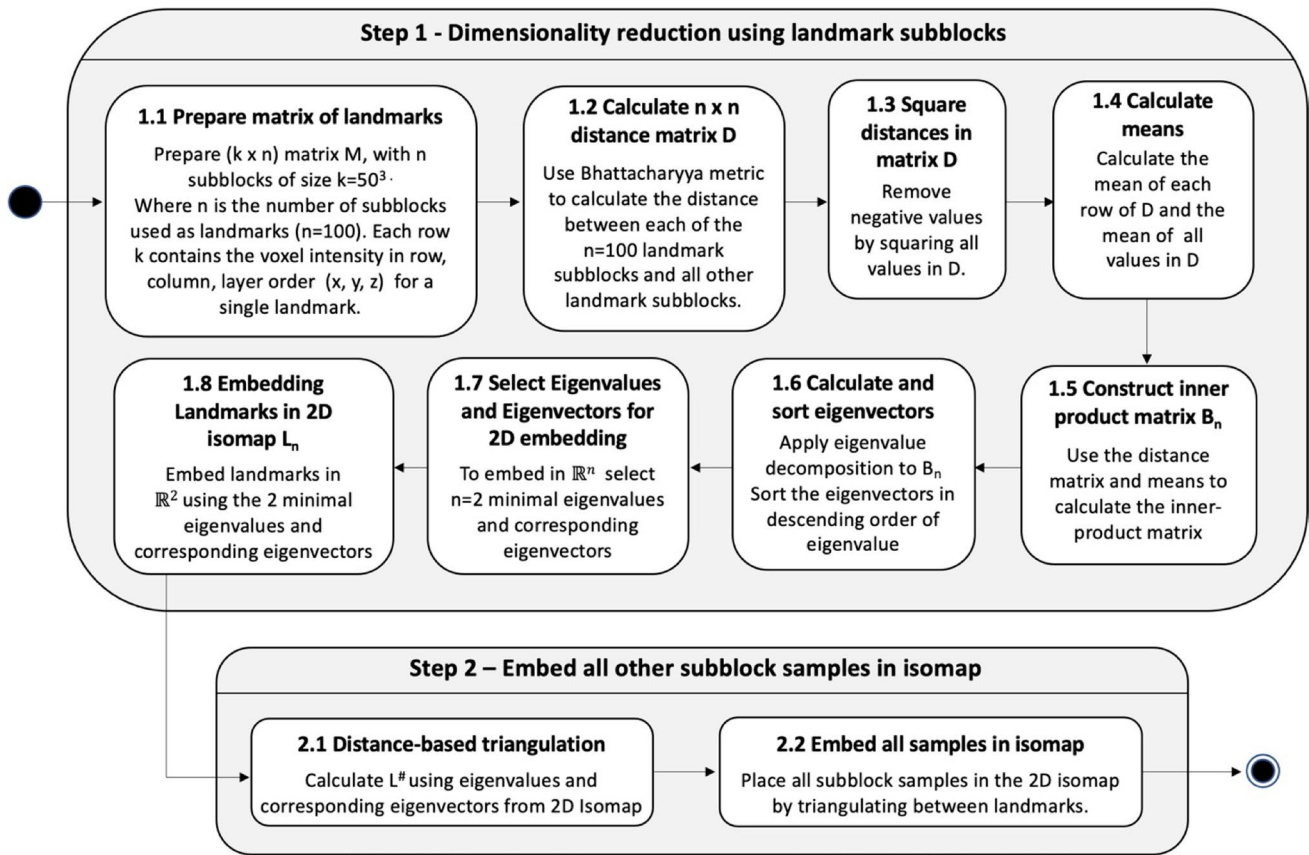


Fig. 6 Two-step algorithm for landmark multidimensional scaling (De Silva and Tenenbaum 2004)

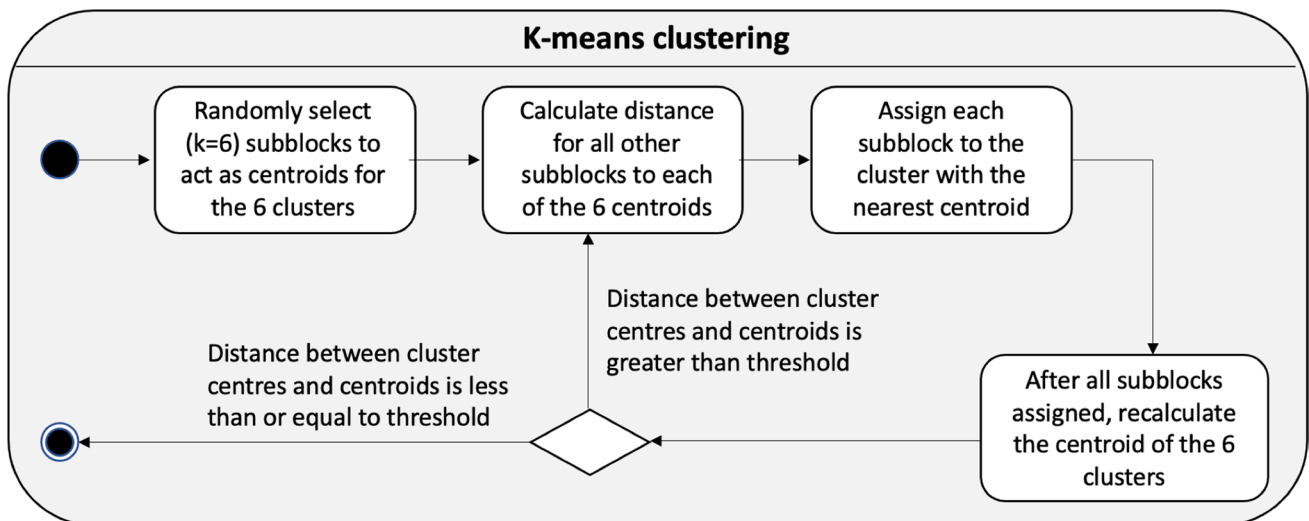
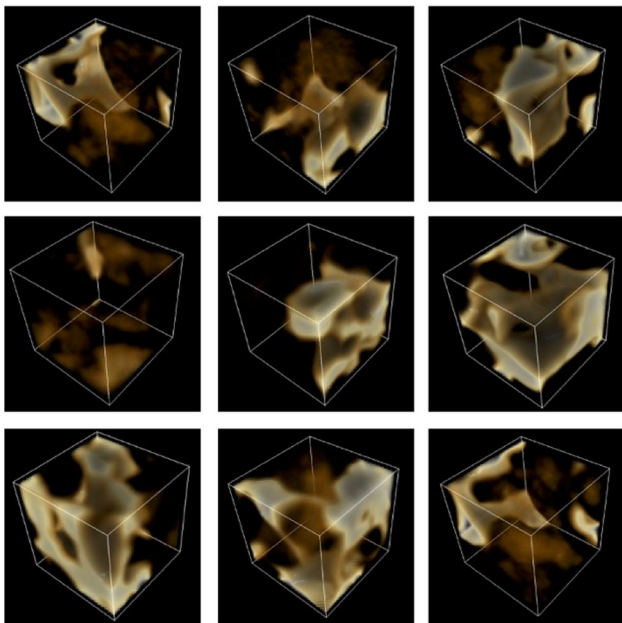
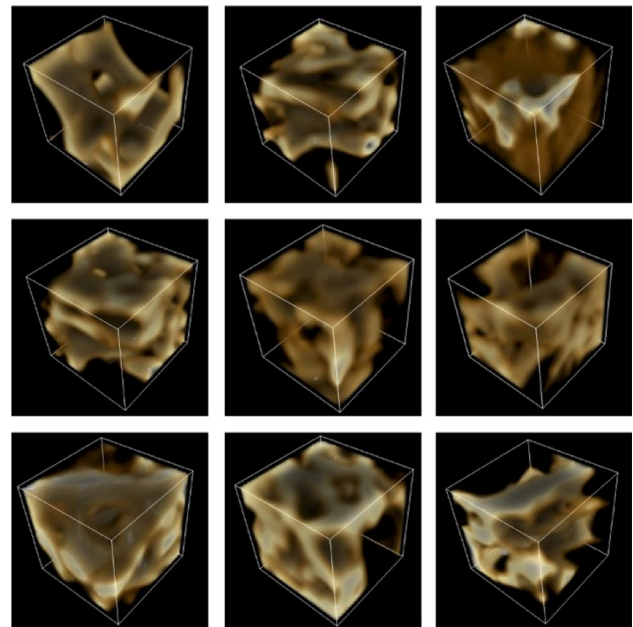


Fig. 7 K-means clustering algorithm

Table 3 Percentage of features in each cluster for the nine cokes analysed

Coke No.	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
1	0.8	4.6	29.9	49.5	12.7	2.5
2	0.5	22.9	42.2	21.5	10.8	2.1
3	1.8	27.2	49.0	19.2	2.7	0.1
4	2.9	22.7	37.8	25.5	9.7	1.4
5	4.4	18.4	60.3	13.9	2.7	0.3
6	2.2	28.5	44.1	14.5	8.9	1.7
7	4.4	25.1	36.7	22.7	10.0	1.1
8	0.4	5.6	48.1	33.7	11.1	1.1
9	8.0	32.7	37.3	16.6	4.5	0.9
Mean	2.8	20.9	42.8	24.1	8.1	1.2

**Fig. 8** Examples of class 1 subblocks**Fig. 9** Examples of class 2 subblocks

supports visualization of volumetric data. Due to the large number of subblocks generated, one thousand per sample, it is not possible to manually consider each block in detail. Rather a random selection of ten classified 50^3 voxel subblocks from each class, and for each sample, were examined by domain experts.

We used 8-bit greyscale CT images, with 255 representing white, and an intensity of 0 corresponding to black. Image intensities in the CT scans correspond to the density of structure so white regions correspond to dense artefacts, while blacker regions represent pores. Examples of typical subblock structure for classes 1–6 are shown in Figs. 8,

9, 10, 11, 12, 13. Note that there is an ordering of classes based on the distribution of voxel intensities and that lower class (1–2) subblocks tend to be dominated by pores. Typically, there is an increase in wall thickness and reduced pore sizes in the higher classes. Inert features are also likely to be seen in higher classes (5–6). The size of these inert features means that they are often large enough to dominate a subblock.

To further examine the distinguishing features of each of the six classes further analysis was carried out in the subblocks by calculating various first and second order statistics for each of the six classes. These approaches are typically used

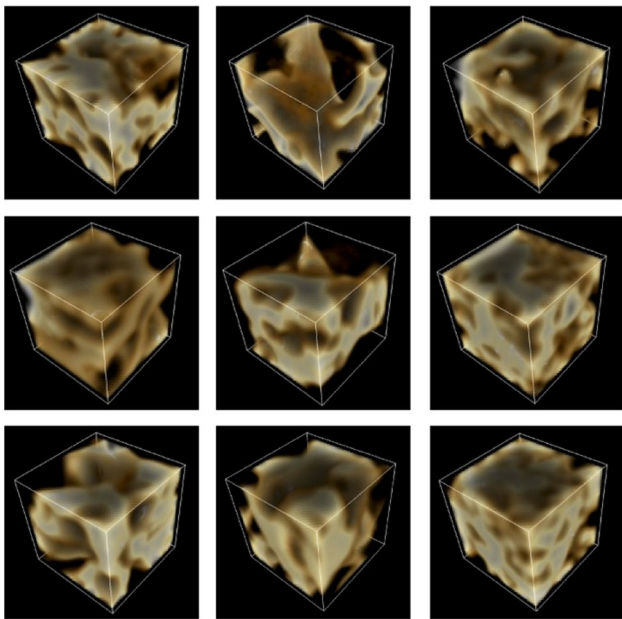


Fig. 10 Examples of class 3 subblocks

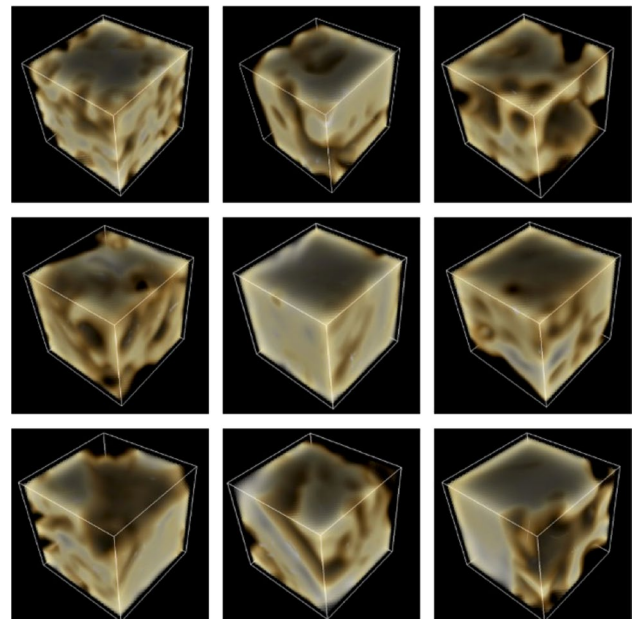


Fig. 12 Examples of class 5 subblocks

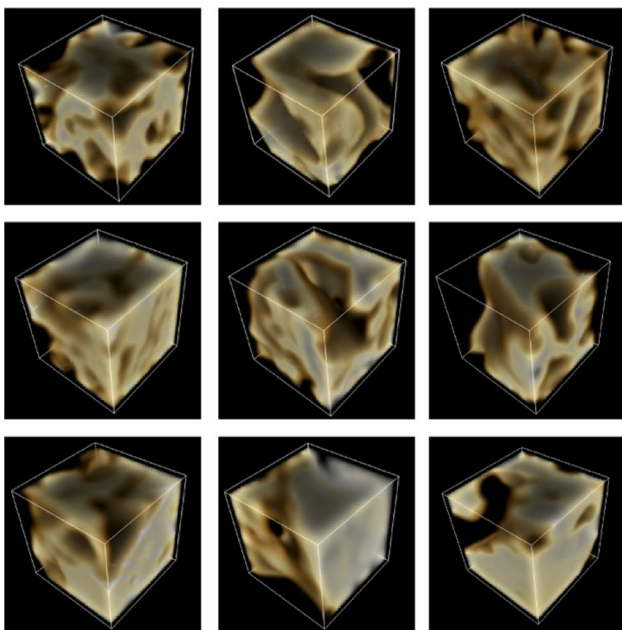


Fig. 11 Examples of class 4 subblocks

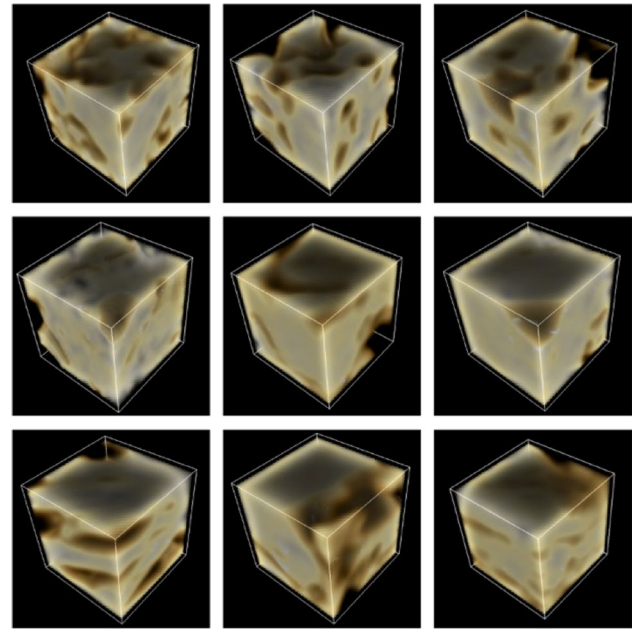


Fig. 13 Examples of class 6 subblocks

for image texture analysis (Materka and Strzelecki 1998; Mailard 2003; Bharati et al. 2004). The combined first order results for all 27 samples are shown in Table 4 and Fig. 14.

The results for all slices in both orientations, were combined for each subblock volume and the average scores for each class are shown in Tables 4, 5 and Fig. 14.

4 Discussion

This project successfully used LMDS with the Bhattacharyya distance metric and k-means clustering to classify 50^3 voxel subblocks of CT images of coke. This approach produced a well-ordered group of classes, with lower classes

Table 4 First order statistics for combined subblocks from each class

Measure		Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Mean intensity	$\mu = \sum_{i=0}^{G-1} ip(i)$	23	39	67	76	96	116
Energy	$E = \sum_{i=0}^{G-1} [p(i)]^2$	0.041	0.026	0.017	0.009	0.010	0.007

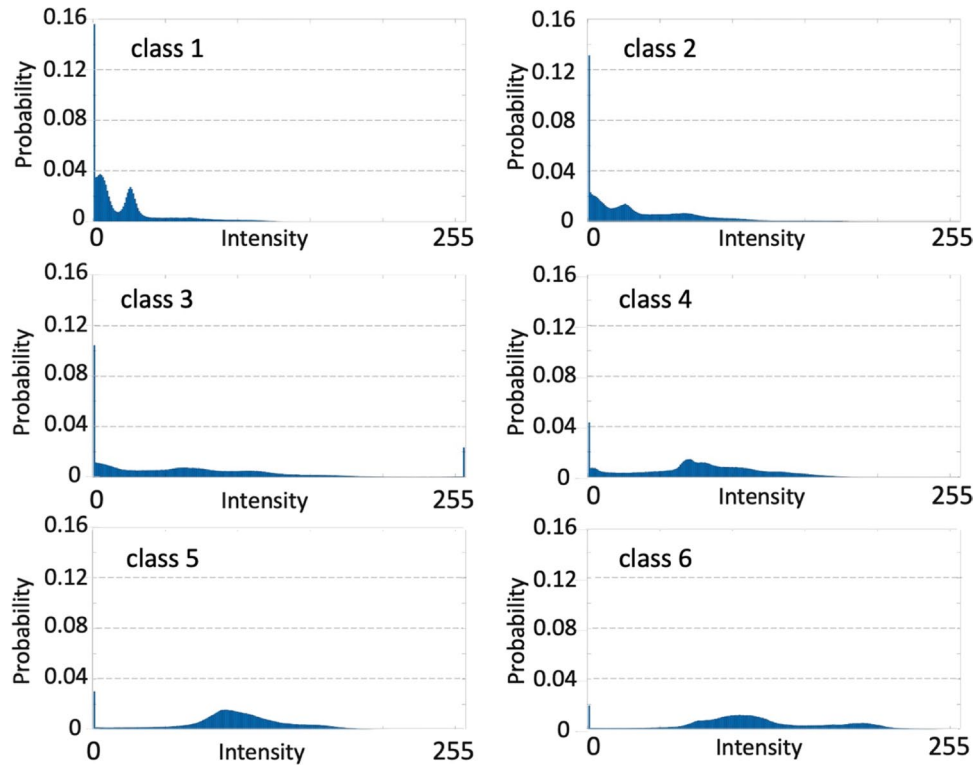


Fig. 14 Probability distribution of voxel intensities for each class of subblocks

Table 5 Second order statistics for combined subblocks from each class. Results were calculated by considering each 2D slice in the subblock volume in two orientations and were required using 8 grey levels

Measure		Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Inertia (contrast) (8 grey levels)	$\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i-j)^2 p(i,j)$	0.07	0.15	0.23	0.17	0.19	0.18
Energy (8 grey levels)	$\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} [p(i,j)]^2$	0.61	0.32	0.16	0.16	0.18	0.15

dominated by porous areas and thin walls. In general, pore size and wall thickness increased from class 1 to class 6. Class 5 and class 6 microstructure tended to be dominated by thick pore walls and inert components. As seen in Fig. 14, the Bhattacharyya metric is also reflected in the final clustering with each class showing different patterns of intensity distribution.

In this project we employed interactive volumetric tools, using 3D visualization to help check the final classification of subblocks. Heuristically, the approach was found to produce an ordered although uneven classification of the subblocks. Class 3 microstructure was the most generic across samples, with 43% of blocks classified in this way. This suggests that this type of microstructure may be a common

feature of the cokes analysed. Both class 2 (21%) and class 4 (24%) were also common to most of the cokes, although variations were found across samples. Class 1 (3%), along with class 5 (8.1%) and class 6 (1.2%) were less common than other classes but may also be more important for differentiating coke quality as class 1 are associated with thin pore walls and large pores, and class 5 and class 6 which tend to include thick pore walls and large inert elements. The distribution of the less common classes of features may for example indicate locations of structural strength or weakness.

It is tempting to consider the distribution of these classes across the nine cokes to see if they provide some indication of differences in the coke performance (Fig. 14). For example, it might be expected that knowing the number of low-density structural points is enough to differentiate one factor of coke strength. Although given the heterogenous nature of coke we might also expect that the number of points in each class would vary between samples, even of the same coke. In fact, the variability across all cokes was found to be statistically significant. Thus, even when comparing samples of the same coke there was a significant difference in the frequency with which different structural features (classes 1–6) occur.

This variability highlights the need to repeat the classification process across many more samples and to potentially increase the size of the sample block. In this work we selected a 500^3 voxel block for analysis as we could automatically extract such a block from the centre of a previously scanned samples. This helped avoid edge effects from the irregularly shaped samples. It might be ideal to consider this when initially scanning samples and allow for a 1000^3 or 2000^3 voxel block for analysis. Although it should also be noted that larger samples may introduce concerns over the processing power required for approaches such as LMDS.

Regardless of the number of classes found for each coke, characterizing coke qualities such as strength may not just be indicated by the presence of weak or strong components. Rather, the way such features are distributed in 3D may be of equal importance. For example, if weak components tend to align they might provide a line or plane of weakness that fails under compressive load. The general approach used in this project is well suited to producing 3D points, each associated with one of the microstructure classes and thus allowing for further study of their 3D distribution patterns.

Since the classes identified in this approach are ordinal in nature, one further benefit of the LMDS approach is that once the landmarks are calculated the class membership of a smaller subblock, at any location in the coke volume, can be estimated by averaging the membership of that smaller block in all overlapping subblocks of the original size.

As described previously, an important element of classification is the ability to group similar blocks of microstructure

by comparing them against one another. In this work the LMDS algorithm relies on a distance metric that calculates the similarity or difference between 50^3 voxel subblocks of coke. Each subblock of coke is reduced to a single vector of length 125,000 (50^3) containing the image intensity at each voxel in the subblock. The distance metric is then used to measure the distance between two such vectors. In this project we used the Bhattacharyya distance metric. Since this metric calculates similarity based on the way image intensities are distributed it provides an isomorphic measure. That is, the sample orientation was found to have limited impact when comparing subblocks of coke.

The decision to adopt an isomorphic measure for microstructure helps to simplify the analysis. However, as microstructure is not necessarily isomorphic an alternative approach might be to consider a different distance metric and repeat the classification using various orientations of the block of coke. At the start of this project, we trialled the Euclidean, Manhattan and Chebyshev metrics but found the Bhattacharyya metrics was the most reliable for classification.

One issue with the LMDS approach in general is that by reducing the spatial dimension, important 3D spatial relationships in the subblocks may be difficult to interpret or even not considered during classification. The 3D shape of microstructure may be quite important. For example, symmetric and regular (isotropic) microtextures react more readily in the presence of carbon dioxide, while flow type, non-symmetric (anisotropic) structures along with coarse mosaic structures show strong resistance to gaseous attack.

An alternative to the LMDS (De Silva and Tenenbaum 2004) approach used in this project is to replace this step with a convolutional autoencoder (Wang et al. 2016). A 3D convolutional autoencoder can be configured to preserve 3D structural cues for use in the classification. Autoencoders have previously been employed for nonlinear dimensionality reduction in several fields where high dimensional data processing and pattern learning are required (Sakurada and Yairi 2014; Beteuhoh et al. 2006; Finn et al. 2016).

One further factor that might need to be considered carefully for classification is the scale at which structural properties impact on coke quality. For example, it may be that thin-walled pore structures are best classified at the scale we have used in this work, by classifying $450^3 \mu\text{m}^3$ blocks. However, the role of larger structures such as inerts may be better characterized at much larger scales. Conversely, the actual binding of inerts into the overall structure may need to be analysed at a much finer scale.

This work does not consider structure formation mechanisms, so no attempt was made to relate physical structure back to chemical mechanisms. The size of the graphitic domains, represented by the microtextural characteristics, is important particularly for reactivity. These are usually

measured by recording the size and extent of the optical anisotropy measured using polarized visible light, however, we can't identify these in the X-ray CT images used in this project. Although these properties are important for coke, our focus is on mechanical properties. Young's modulus in particular has been shown not to vary greatly for the different fused carbon phases (Andriopoulos et al. 2003). Therefore, in this early work we treat the solid phase as having a single set of properties and then it is the balance between pores and wall thickness that is key. We also note that inerts do have some different physical properties from the fused phases and this needs to be addressed in follow up work.

In terms of first order statistics (Tables 3, 4, 5) we see an increase in mean intensity from class 1 to class 6. This might be expected with the Bhattacharyya metric which itself relies on mean image intensity and variance when comparing subblocks of coke. First order energy is a measure of the localized change in an image and as we might expect the higher number of thin walls in class 1 results in a higher energy measure. This reduces for class 2 and class 3 and becomes consistently low for classes 3, 4 and 5 where the subblocks are more consistent in structure. These results suggest both the mean intensity of a subblock, and the energy may help distinguish some classes of subblocks.

In terms of second order image statistics (Tables 3, 4, 5) both second order energy and contrast measures also seem to reflect the presence of more convoluted edges associated with pore walls in class 1 and class 2. Some caution needs to be applied here as we have reduced the number of grey levels in these images to calculate the second order statistics. This may remove some of the more subtle intensity variations associated with class 6 structure such as inerts.

5 Conclusions

Motivated by the need to help develop further fundamental understanding of the processes that underpin the way coke properties are derived from parent coals, this study investigated a novel approach for automatically classifying small $450^3 \mu\text{m}^3$ subblocks of coke microstructure taken from 3D CT images.

The key outcomes of the work were:

- (1) The dimensional reduction approach known as LMDS using the Bhattacharyya distance metric and k-means clustering was able to classify six different classes of microstructure automatically and reliably in 27 coke samples.
- (2) The approach produced well-distributed, well-ordered clusters of microstructures and this classification was preserved under changes in block orientation. Class 1

subblocks tend to be dominated by pores and thin wall structure. Class 2 blocks have less pores and slightly thicker walls. The amount of wall structure continues to increase through classes 4–6, with thicker walls and smaller pores identified. Inerts begin to be present in class 4 and increase in frequency through class 5 and class 6 and often dominate the class 6 subblocks.

- (3) All 27 coke samples and the 9 different cokes showed heterogeneous distributions of microstructure, suggesting further samples need to be analysed to draw precise conclusions. Availability of 3D CT scans may be a limiting factor in this analysis.
- (4) All cokes showed the dominant presence of class 3 features (43%) suggested a common base to the cokes. In general, it may be that less commonly identified microstructure such as class 1 (3%), along with class 5 (8.1%) and class 6 (1.2%) are important for differentiating coke quality as class 1 are associated with thin pore walls and large pores, and class 5 and class 6 show thick pore walls and often contain large inert elements.
- (5) A key benefit of the approach used in this work is that it allows for automatic identification of types of microstructures at 3D locations that naturally allow for the study of the way these features are distributed in 3D. However, it should be emphasized that physical properties at different scales may need to be considered. For example, the binding of inerts cannot be determined from the resolution of scans in this work. It is also not possible to identify 3D carbon structures associated with the identified classes.
- (6) Some of the processes used in this project might be further improved to increase the validity of results. This includes extending the number of samples used in the characterization to counter the natural variations found in samples. It also suggests the evaluation of alternative techniques, such as autoencoding, that more transparently identify 3D structure for use in classification. Although the LMDS algorithm used in this project classifies 3D blocks of microstructure, these blocks need to be reduced to a single dimension for clustering and this introduces a level of abstraction that potentially hinders interpretation of outcomes.

Acknowledgements This research project was supported by funding from the Australian Coal Industry's Research Program (ACARP)—Grant Number C29073. This research was undertaken on the Imaging and Medical Beamline at the Australian Synchrotron, part of ANSTO. The work was carried out with the assistance of the Drishti/Drishti Prayog visualization software developed by the National Computational Infrastructure's VizLab.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source,

provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Andriopoulos N, Loo CE, Dukino R, McGuire S (2003) Micro-properties of Australian coking coals. *ISIJ Int* 43(10):1528–1537. <https://doi.org/10.2355/isijinternational.43.1528>
- Betechuoh BL, Marwala T, Tettey T (2006) Autoencoder networks for HIV classification. *Curr Sci* 91(11):1467–1473
- Bharati MH, Liu JJ, MacGregor JF (2004) Image texture analysis: methods and comparisons. *Chemom Intell Lab Syst* 72(2004):57–71
- Chalup SK, Clement R, Marshall J, Tucker C, Ostwald MJ (2007) Representations of streetscape perceptions through manifold learning in the space of hough arrays. In: Proceedings of the 2007 IEEE symposium on artificial life, IEEE.
- Chen P, Wang P, Han Y, Zhang Z (2014a) Automatic coke microstructures recognition using image segmentation. *Int J Comput Appl Technol* 50(1–2):51–60
- Chen Y, Pfeffer S, Fernández JJ, Sorzano COS, Förster F (2014b) Autofocused 3D classification of cryoelectron subtomograms. *Structure* 22(10):1528–1537
- Chen Y, Lee S, Tahmasebi A, Bai J, Mahoney M, Yu J (2020) A review of the state-of-the-art research on carbon structure evolution during the coking process: From plastic layer chemistry to 3D carbon structure establishment. *Fuel* 5(271):117657. <https://doi.org/10.1016/j.fuel.2020.117657>
- Cox TF, Cox MAA (1994) *Multidimensional scaling*. Chapman and Hall, London
- Finn C, Tan XY, Duan Y, Darrell T, Levine S, Abbeel P (2016) Deep spatial autoencoders for visuomotor learning. In: Okam A (ed) IEEE 2016 international conference on robotics and automation (ICRA). Stockholm, Sweden, pp 512–519
- Ghosh B, Sahoo BK, Chakraborty B et al (2018) Influence of coke structure on coke quality using image analysis method. *Int J Coal Sci Technol* 5:473–485. <https://doi.org/10.1007/s40789-018-0227-0>
- Goudail F, Réfrégier P, Delyon G (2004) Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images. *J Opt Soc Am A* 21:1231–1240
- Hartigan JA, Wong MA (1979) Algorithm AS 136: A k-means clustering algorithm. *J R Stat Soc Ser C* 28(1):100–108
- Kamnitsas K et al (2017) Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 36:61–78
- Kruskal J, Wish M (1978) *Multidimensional scaling*. Sage, Beverly Hills, CA
- Limaye A (2012) Driшти: a volume exploration and presentation tool. In: Stock SR (ed) *Developments in X-ray tomography VIII*, Proceedings of SPIE vol 8506. <https://doi.org/10.1117/12.935640>
- Lomas H, Jenkins DR, Mahoney MR, Pearce R, Roest R, Steel K, Mayo S (2017) Examining mechanisms of metallurgical coke fracture using micro-CT imaging and analysis. *Fuel Process Technol* 155:183–190
- Magdalinos P, Doukeridis C, Vazirgiannis M (2011) Enhancing clustering quality through landmark-based dimensionality reduction. *ACM Trans Knowl Discov Data* 5(2):1–44
- Maillard P (2003) Comparing texture analysis methods through classification. *Photogramm Eng Remote Sens* 69(4):357–367
- Materka A, Strzelecki M (1998) *Texture analysis methods – a review*. Technical. University of Lodz, Institute of Electronics, COST B11 report, Brussels.
- North L, Blackmore K, Nesbitt K, Mahoney MR (2018) Methods of coke quality prediction: a review. *Fuel* 219:426–445. <https://doi.org/10.1016/j.fuel.2018.01.090>
- Pandit S, Gupta S (2011) A comparative study on distance measuring approaches for clustering. *Int J Res Comput Sci* 2(1):29–31. <https://doi.org/10.7815/ijorcs.21.2011.011>
- Qi CR, Su H, Niessner M, Dai A, Yan M, Guibas LJ (2016) Volumetric and multi-view cnns for object classification on 3d data. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) 5648–5656
- Rantitsch G, Bhattacharyya A, Schenk J, Lünsdorf NK (2014) Assessing the quality of metallurgical coke by Raman spectroscopy. *Int J Coal Geol* 130:1–7
- Riley JT (2007) Routine coal and coke analysis: collection, interpretation, and use of analytical data. *Micro & Nano Letters* 84–92.
- Sakurada M, Yairi T (2014) Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Rahman A, Deng J, Li J (eds) *Proceedings of the MLSDA 2014 2Nd workshop on machine learning for sensory data analysis*, ACM, New York, pp 4–11
- Shepard RN (1980) Multidimensional scaling, tree-fitting, and clustering. *Science* 210:390–398
- Shirkhorshidi AS, Aghabozorgi S, Wah TY (2015) A comparison study on similarity and dissimilarity measures in clustering continuous data. *PLoS ONE* 10(12):e0144059. <https://doi.org/10.1371/journal.pone.0144059>
- De Silva V, Tenenbaum JB (2004) *Sparse multidimensional scaling using landmark points*. Technical report (Vol. 120), Stanford University.
- Tenenbaum JB, de Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. *Science* 290:2319–2323
- Tiwari HP, Haldar SK, Roy A, Bhattacharjee A, Mishra P (2015) Data mining—new perspectives on predicting coke quality in recovery stamp charged coke making process. *Metall Res Technol* 112(6):603
- Tsafnat N, Tsafnat G, Jones AS (2008) Micro-finite element modelling of coke blends using X-ray microtomography. *Fuel* 87(13–14):2983–2987
- Wang Y, Yao H, Zhao S (2016) Autoencoder based dimensionality reduction. *Neurocomputing* 184:232–242
- Wu J, Xie M, Zeng Y, Liang C, Che Y, Zhu J, Wei Q, Guo T, Jiang Z, Liu F, Hu, D. (2022). Microstructure analysis and mineral phase transformation of coke in blast furnace. *Ironmaking & Steelmaking*, 1–10.
- Zhang R, Tao J (2017) Data-driven modeling using improved multi-objective optimization based neural network for coke furnace system. *IEEE Trans Ind Electron* 64(4):3147–3155. <https://doi.org/10.1109/TIE.2016.2645498>
- Zheng H, Xu R, Zhang J, Daghighaleh O, Schenk J, Li C, Wang W (2021) A comprehensive review of characterization methods for metallurgical coke structures. *Materials* 15(1):174
- Zou LJ, Zheng J, Miao C, Mckeown MJ, Wang ZJ (2017) 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. *IEEE Access* 5:23626–23636. <https://doi.org/10.1109/ACCESS.2017.2762703>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.