Wear debris from total joint replacements:

Eckold, David; Dearn, Karl; Shepherd, Duncan

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Eckold, D, Dearn, K & Shepherd, D 2016, 'Wear debris from total joint replacements: evaluation of automated categorisation by scale-invariant feature transforms', *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*. https://doi.org/10.1080/21681163.2016.1230075, https://doi.org/10.1080/21681163.2016.1230075

Wear debris from total joint replacements: evaluation of automated categorisation by scale-invariant feature transforms

David G. Eckold, Karl D. Dearn & Duncan E. T. Shepherd

Submit your article to this journal

View related articles

View Crossmark data

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Wear debris from total joint replacements: evaluation of automated categorisation by scale-invariant feature transforms

David G. Eckold, Karl D. Dearn (ID) and Duncan E. T. Shepherd

Department of Mechanical Engineering, School of Engineering, University of Birmingham, Birmingham, UK

ABSTRACT

Wear debris is a crucial factor in determining the lifespan of a total joint replacement. Not only do particulates between a bearing surfaces increase wear rates through third-body abrasion, but immune reactions can lead to inflammation and osteolysis. In this paper, the use of computer vision to analyse and classify scanning electron microscope images of debris was investigated. UHMWPE debris was generated using an *in vitro* simulator or a linear tribometer, images were analysed using scale invariant feature transforms and a support vector machine classifier. The accuracy was 77.6% with a receiver operating characteristic area under curve of 92%.

1. Introduction

The characteristics of wear debris can be related to the wear mechanics of a bearing system (Anderson 1982). These characteristics have particular importance in joint replacement implants as not only do they indicate the wear regime (Kumar et al. 2013), but they also influence immune reactions of the patient (Ingham & Fisher 2000; Hallab & Jacobs 2009).

The wear of joint replacement implants has been studied in detail (Punt et al. 2011; Xin et al. 2013; Neukamp et al. 2014; Moghadas et al. 2015) and the debris from these studies is often examined using scanning electron microscopy (Nine et al. 2014), as it has excellent imaging properties at the magnification ranges required to produce highly detailed micrographs of debris. While it is easy to characterise debris on simple metrics such as equivalent circle diameter (ECD), aspect ratio (AR) and roundness, the task of categorising the contents of any image in qualitative terms i.e. morphology, remains a significant challenge.

Various studies have attempted to create bespoke wear debris recognition algorithms, making use of the various properties of the debris particles (Kirk et al. 1995; Zhang et al. 1997; Podsiadlo & Stachowiak 2000). For example, partition-iterated function systems (PIFS) (Stachowiak & Podsiadlo 1999) and wavelet (Chen et al. 2006) methods take advantage of the fractal nature of wear debris, i.e. there are recurring features at various scales. These methods have been shown to be extremely powerful, and have high success in terms of accuracy.

At the time of development of these classifiers, the field of computer vision was nascent and emerging methods and algorithms were performing relatively poorly (Boiman et al. 2008).

However, advancement in the field has occurred quickly throughout the mid–late 2000s (Everingham et al. 2010).

There are two key areas of improvement within this field: that is the classifier, and the descriptor quantiser. The classifier is the mechanism for classifying objects based on the data fed to them; these can be learning based, or non-parametric based. Common classifiers are support vector machine (SVM) (Cortes & Vapnik 1995; Zhang et al. 2006), decision trees (Bosch et al. 2007) and nearest neighbour (Boiman et al. 2008) based classifiers. The descriptor quantiser describes the algorithms a computer uses to interpret an image and generates quantitative data about an image that describes what is pictured within. This can be as simple as colour histograms, or object size/aspect ratio, to more complex properties such as fractal dimensions, geometric-blur (Zhang et al. 2006) or the descriptor quantiser used in this paper – scale-invariant feature transforms (SIFT).

A comparison of various descriptor algorithms was performed by Mikolajczyk and Schmid (2005); their results showed SIFT performed well against comparable methods for locating and describing key points. This algorithm combined with the SVM classifier performed well in the Pascal Visual Object Classes Challenge 2012 (Everingham et al. 2015). Additionally, all the highest scoring methods for classification and detection used an SVM for classification (Everingham et al. 2015).

Wear debris analysis is of particular interest for orthopaedic applications, as wear debris has been linked to the immune reactions leading to osteolysis (Purdue et al. 2006). This reaction to wear debris has been shown to be affected by size and shape (Green et al. 1998, 2000; Yang et al. 2002). Despite continued work on the analysis of wear debris in biomedical engineering, where debris from *in vivo* and *in vitro* orthopaedic implant studies

---

CONTACT Karl D. Dearn ✉ k.d.dearn@bham.ac.uk

is regularly characterised (Kumar et al. 2013; ; Nine et al. 2014; Hongtao et al. 2011; Eckold et al. 2015; Saikko et al. 2015), there have been few advancements recently on moving beyond simple size and shape attributes. Using more sophisticated tools for the analysis of wear debris, new insights into which debris and wear regime are more likely to lead to failure could be made. New orthopaedic implants could then be designed to avoid this debris.

This paper describes an implementation of, and the viability of using, an open-source, but well-regarded and robust generic object recognition algorithm for the use of wear debris analysis from scanning electron microscope (SEM) micrographs. The aim is to introduce methodologies from outside disciplines with greater experience in computer vision, thereby allowing biomedical engineers and tribologists the opportunity to analyse SEM images without the need to reinvent tools found elsewhere. By removing the obstacle of creating a program that can recognise debris, greater comparison can be made between papers on the subject of wear debris analysis. The algorithm used in this paper is known as SIFT, and was invented by Lowe (2004); the implementation is known as VLFeat from the Oxford Vision Laboratory (Vedaldi & Fulkerson 2010b, 2010c).

2. Methods

2.1. Debris generation

Wear debris was generated, isolated and then imaged using SEM. The material used to generate the debris was ultra-high molecular weight polyethylene (UHMWPE) and it was generated using two different methods: (1) a Bose ElectroForce Spinal Disc Fatigue/Wear system (SD-F/W) (Bose Corp., ElectroForce Systems Group, Eden Prairie, Minnesota, USA); or (2) a high-frequency reciprocating rig (HFRR) (PCS Instruments, London, United Kingdom).

The *in vitro* UHMWPE debris was created in a study by Moghadas et al. (2015) using the Bose SD-F/W to wear test a Charité total disc replacement (TDR) implant. This device has two cobalt–chrome–molybdenum' (CoCrMo) concave end-plates and a central convex UHMWPE core (Figure 1). The Charité was loaded' sinusoidally at 2 Hz between 500 and 2000 N, according to ISO standard 18192-1:2011 (British Standards Institution 2011b). Rotations were between 6° and 3° in flexion/extension, and between 2° and 2° in both lateral bend and rotation (Figure 2). The lubricant in the Bose SD-F/W was bovine serum albumin (BSA) (30 gl⁻¹ protein content) (Sera Laboratories Int, West Sussex, United Kingdom).

The HFRR is a linear reciprocating motion ball on disc tribometer. The test used to generate debris used a 6-mm diameter ANSI E52100 steel ball (PCS Instruments, London, United Kingdom) on GUR 1120 UHMWPE discs (Orthoplastics, Lancaster, UK). The tribometer used a 2-mm stroke length. In one test, the ball was roughened with P400 grit wet and dry paper (3 M, St. Paul, Minnesota, USA) to Ra = 0.5 µm to increase the amount of abrasive debris generated; in this test, a 20 Hz frequency was used. To generate the adhesive wear, a smooth ball was used (Ra = 0.05 µm) at a frequency of 25 Hz. The lubricant used in the HFRR was ultrapure deionised water (resistivity > 18 MΩ cm, inorganic content < 2 ppb), and was selected to simplify the tribological mechanisms involved in generating and controlling

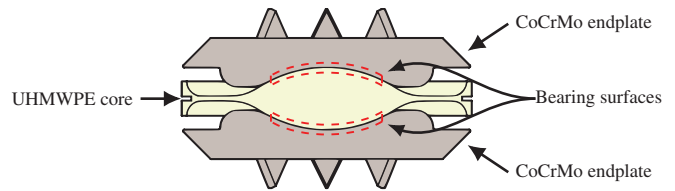**Figure 1.** Illustration of the cross section of a Charité implant.
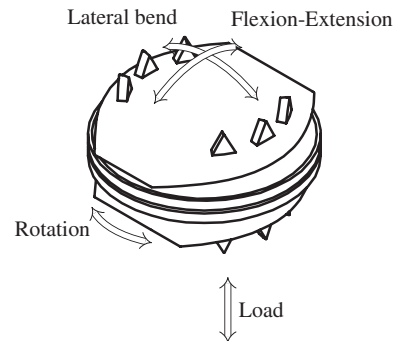

Figure 2. Directions of rotations exerted on the Charité implant.

wear debris morphology. The use of ultrapure water simplified the classification of debris particles by avoiding large-scale lubricant-born contaminants.

2.2. Debris isolation

The wear debris was collected from the lubricants through vacuum filtration. The bovine serum was first digested to remove attached proteins and other biological contaminants found in bovine serum using the hydrochloric acid (HCl) method outlined in BS ISO 17853:2011 (British Standards Institution 2011a, p. 9). This method uses hydrochloric acid at 50 °C to digest the biological content in BSA; it was then diluted with methanol to reduce the viscosity for filtering.

Subsequent to digestion, the debris containing fluids was filtered through 0.1 µm nuclepore filters (Whatman International Ltd, Maidstone, United Kingdom) in a vacuum filtration system, and then was mounted on an SEM aluminium stub and sputter coated in gold for 60 s at 30 mA using an Agar automatic sputter coater (Agar Scientific, Elektron Technology UK Ltd, Essex, United Kingdom).

2.3. SEM imaging

SEM images were taken on either a Jeol 7000F FEG-SEM (Jeol Ltd., Tokyo, Japan) or an FEI Duelbeam FIB-SEM (FEI, Hillsboro, Oregon, USA). Electron voltages were 10 kV unless the debris started to suffer from beam damage (e.g. swelling or cracking), in which case the beam voltage was lowered to 5 kV. Secondary electron images were collected from the typical Everhart-Thornley detector on the Jeol 7000F; however, images generated by the FEI system used the 'through lens detector' on 'ultra-high resolution' mode.

Images were focused and astigmatism was corrected for, the magnification was chosen to achieve a full frame image of the

particle. A long scan for 26 s was taken without averaging to produce a clear, low-noise image that was saved as a TIF file.

2.4. Image processing & machine learning

The images were sorted by humans into classes based on the following criteria:

Adhesion – wear debris generated by the HFRR in the smooth ball, higher frequency setting (Figure 3(a) and (b)).

Chip – wear debris generated by the HFRR in the rough ball setting (Figure 3(c) and (d)).

Fibril – wear debris that had formed into long, thin fibrils (Figure 4(a) and (b)).

Large Sphere – wear debris that was of > 5 μm in diameter and appeared to be of a spherical shape (Figure 4(c) and (d)).

Flake – wear debris that appeared flat and sheet-like (Figure 4(e) and (f)).

Images were processed using the MATLAB image processing toolbox to remove the background. This ensured that features on the filter that may be common between different images, did not cause erroneous positive matches that were not based on the particle morphology. The background was removed from the image using edge-based image segmentation; the outline of the particle was found using 'Canny' edge detection (Canny 1986). Examples of wear debris once the background had been stripped are shown in Figures 3 and 4.

SIFT descriptors were generated for the images, and they were in the form of 128-dimension vectors that described the gradient vector of intensity within a 4×4 grid. Five images of each class were selected pseudo-randomly to be training images. To increase computational efficiency, rather than comparing the descriptors of unknown images against the entirety of the descriptor data of the training images, clusters of similar descriptors were found and averaged. The individual means of these clusters formed the 'words' used in the solver's vocabulary and this is known as a 'bag of words' method. An example using four clusters of random

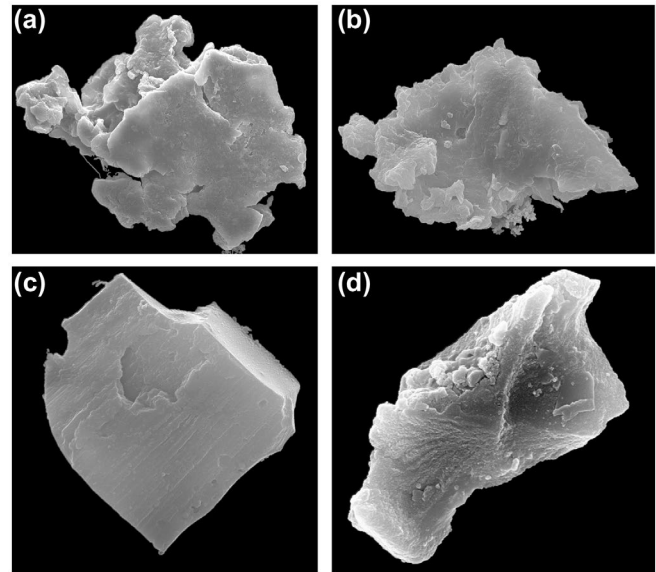**Figure 3.** (a), (b) Example training images for the adhesion class. (c), (d) Examples of debris of the chip class.
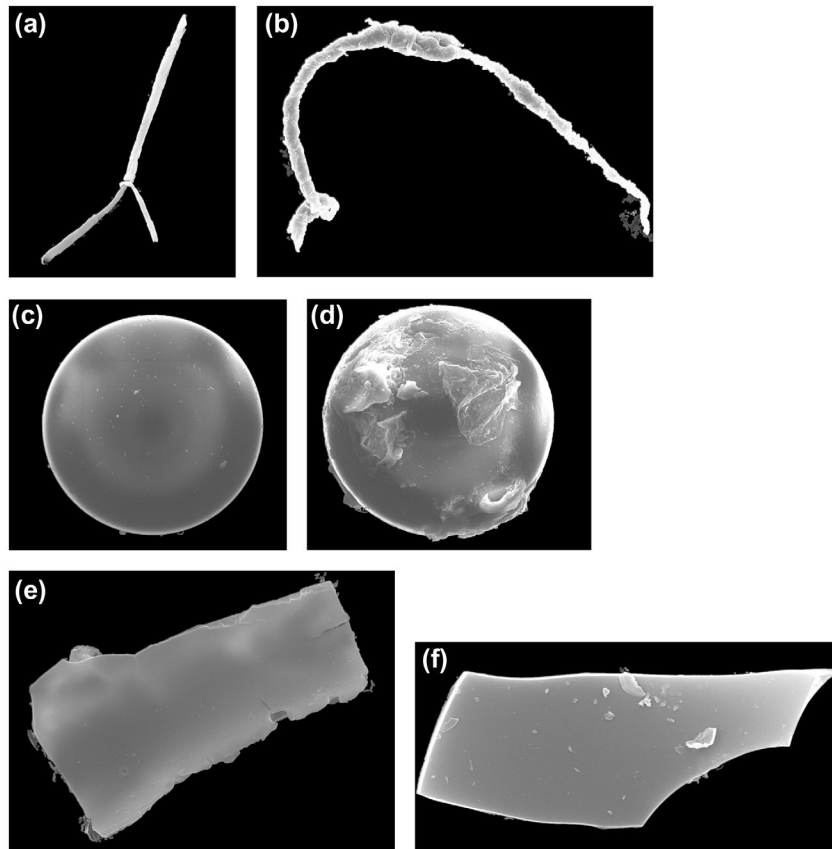

Figure 4. (a), (b) Examples of debris of the fibril class, (c)–(d) large spherical debris and (e), (f) are sheets/flake debris.

coordinate data is shown in Figure 5, the blue dots are the cluster means, and the lines are the Voronoi polygons that divide the clusters.

To generate the 'words' used to describe the particles, all the descriptors were concatenated, and the means of the clusters of descriptors were found. The mean value for each cluster was found using a *k*-means algorithm (Hartigan & Wong 1978), where *k* is the number of 'words'. An amount of 'words' were chosen based on a preliminary study, measuring the accuracy of the solver and computational time. As shown in Figure 6, the complexity was linear with respect to the number of 'words', but the accuracy did not improve beyond 600 'words'.

For rapid nearest neighbour searches, i.e. finding the closest matching 'word' for the descriptor in Euclidean space, the vocabulary was indexed by generating a *k-d* tree (Bentley 1975). This is an efficient way to find the 'word' with the shortest orthogonal distance between descriptor vectors.

To make comparisons between images of how far their descriptors differ from the 'words' efficiently, the closest match (which did not exceed a threshold) was tallied and a histogram of how many descriptors matched what 'word' was computed. The histogram can be seen as a compact précis of the descriptor data.

The SVM is a binary solver – it only categorises something as belonging to a class, or that it does not. To train the SVM, it requires both the histograms of the training image which belong to a class and all of the training histograms that do not. The SVM used is a linear SVM; however, a homogenous kernel map is used to approximate the χ^2-kernel (Vedaldi & Zisserman 2012). This reduces the computation time to train the SVM dramatically while being indistinguishable in performance (Vedaldi & Zisserman 2012)

2.5. Performance analysis

For each class listed in 2.4, the SVM is trained on a random subsection of all the images; the SVM is taught what is a member of a given class vs. all other classes. Once trained, all non-training images are scored against each given class.

The robustness of the method was tested with repeated random subsampling cross validation. By varying the seed used in the pseudo-random number generator, different images were selected for training and testing, the accuracy of the image classification algorithm was assessed using different combinations of training and test images. The process of training and assessing images was repeated five times with different seeds and averaging the results; the number of matches for each class was then found. For each random permutation of images, half were used for training, and half for testing.

The number of matches found was subdivided into which class the image truly belonged , thereby generating a confusion matrix of true and false positives.

3. Results

3.1. Accuracy

Table 1 shows the overall accuracy of the SVM for all the classes, the average accuracy ± the standard deviation was 77.60 ± 4.56%. The accuracy was found by taking the mean of
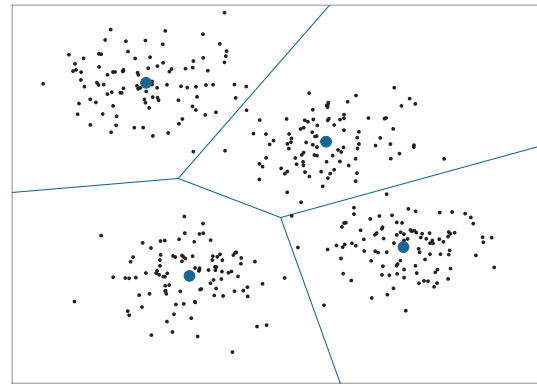
Figure 5. Example of *k*-means clustering using four clusters of normally distributed random coordinates (black dots). The blue dots are the cluster means and the lines are the Voronoi polygons which partition the clusters.
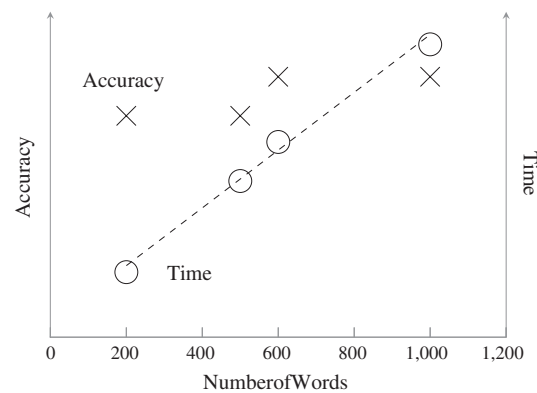
Figure 6. Accuracy and time against number of 'words' in vocabulary. Note the time increases linearly with the number of 'words', but accuracy was constant except for the step increase at 600 'words'.

Table 1. Accuracy of SVM. Average accuracy = 77.6 ± 4.6%.

Random Seed	1	2	3	4	5
Accuracy	80%	76%	76%	72%	84%

Table 2. Confusion matrix of SVM. Rows indicate the particle class the SVM has been trained to find and the columns indicate per cent of particles determined to be a match.

	Large sphere	Adhesion	Fibril	Chip	Sheet
Large sphere	100	0	0	0	0
Adhesion	0	100	0	0	0
Fibril	0	20	68	0	12
Chip	16	20	0	64	0
Sheet	4	0	16	24	56

the true positives (the diagonal of the confusion matrix (Table 2)) of the solver.

3.2. Confusion matrix

Table 2 shows the confusion matrix; this shows the percentage of images found to be a match by the SVM. The rows show which class of debris the SVM was classifying , and the columns show

how many matches were found from each class the image actually belongs to.

3.3. Receiver operating characteristic curves

Figure 7 show the Receiver operating characteristic (ROC) curves for each class of debris. The ROC curves shown are a measure of the true positive rate (recall) against true negative rate at different discrimination thresholds. A good classifier will have both a high rate of true positives and true negatives (the area under curve [AUC] will approach 100%). Random chance is shown with a dashed red line and the performance of the classifier is the solid blue line.

3.4. Timing

Using a 64-bit 2.2 GHz Intel Sandy Bridge E5-2660 CPU with 32 GB available memory, computing the SIFT descriptors and frames took 1.0016 s per image. The MATLAB version was R2015a and the implementation of the SIFT algorithm is from VLFeat 0.9.20.
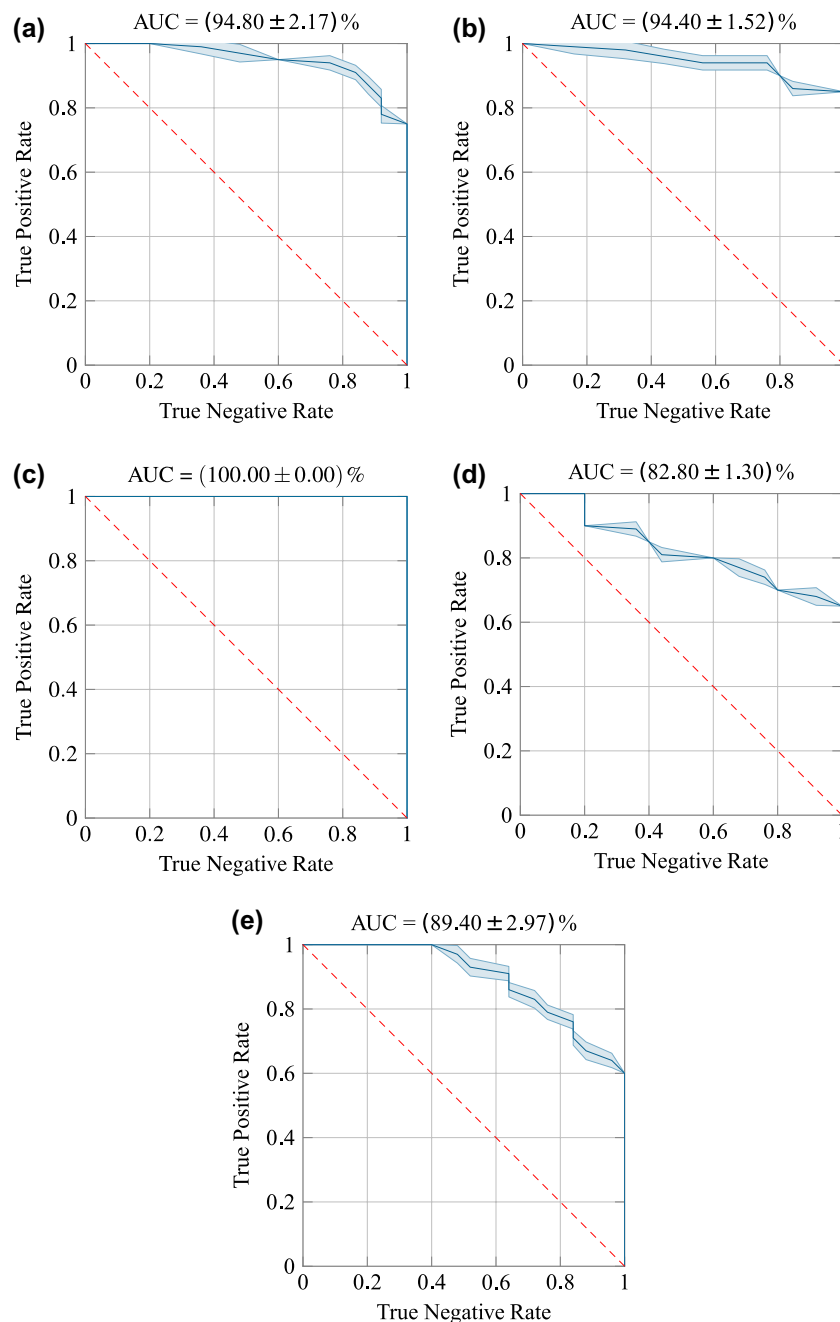
Figure 7. ROC curves of the SVM for each class. (a) ROC curve for fibril debris classification, (b) ROC curve for adhesive particle classification. The ROC curve for spherical particles is (c) and the ROC curve for chips is (d). (e) ROC curve for sheets/flakes. The average AUC of the ROC curves was $92.28 \pm 6.49\%$. The dashed red line (–) is a ROC curve or random chance, the blue line (–) is the mean ROC curve and the shaded area is the standard deviation of the mean curve. AUC = area under curve.

4. Discussion

This study has examined the suitability of SIFT and a parametric classification algorithm typically used for general computer vision, applied specifically for the use of wear debris analysis. It assesses the overall accuracy, as well as the response within each class of wear debris.

The use of image analysis and machine learning for the automated characterisation of debris greatly speeds up the analysis of large quantities of SEM images. It has been shown that the use of general computer vision techniques is applicable for examining micrographs of wear debris when given training images of different debris morphologies, and is comparable to bespoke methodologies (Stachowiak et al. 2008; Kumar et al. 2013). While computers have long been used to do basic analysis for finding size, AR and roundness of particles, the recognition of what an image contains remains a challenging problem in all fields of computer science. Using more general machine learning and image analysis tools than those used in previous studies of debris analysis, the breakthroughs discovered outside of the field of tribology can be re-purposed for examining wear debris with greater accuracy and with fewer inefficiencies attempting to recreate redundant methods.

Wear debris analysis commonly suffers from the subjective nature of interpreting images; by training a computer vision algorithm using debris either of pronounced class, or generated in a sterile environment that greatly favours certain wear regimes, characterising the debris with known confidence levels is possible.

4.1. Accuracy

The overall accuracy of this method, shows a high rate of classification accuracy, correctly identifying debris 77.6% of the time. An accuracy of 77.6% was in line with the capabilities of similar methods using the same descriptor generation and classifier when performed on the Caltech-101 and the PASCAL VOC 2007 classification challenge (Vedaldi & Fulkerson 2010a; Chatfield et al. 2011). Chatfield et al. (2011) found the accuracy of various computer vision methods, including the method presented in this paper, was between 72.3 and 77.78% (SIFT scored 73.77 ± 0.70) on the Caltech-101. Therefore, with no loss in accuracy, the SIFT-based method can be adapted for wear debris recognition. It was found that the accuracy does vary between different random seeds, implying that the quality of training images has some impact on the accuracy of the SVM. The effect of 77.6% accuracy on the reported distributions of wear debris classes is dependent on the prior distribution of each class (Lee 2012). However, the particle classes that take up the majority share will still be reported as such.

4.2. Confusion matrix

From the confusion matrix, it is apparent that a SIFT-based method suffers when a wear particle does not have strongly contrasted features. This is shown by the SVM achieving good accuracy (100%) for 'feature rich' debris, i.e. adhesion (Figure 3(a), (b)) and large spheres (Figure 4(c), (d)), and reasonable accuracy for fibrils (Figure 4(a), (b)) and chips (Figure 3(c), (d)) (64–68%).

However, for 'feature poor' debris like sheet/flakes, (Figure 4(e), (f)) the classifier accuracy was <60% (although still better than chance). While it may appear that large spheres should be 'feature poor', the SIFT algorithm also uses the particles perimeter as a feature, where the circular outline is unique to spherical particles. Considering only a single descriptor type was used for characterisation, the scores demonstrated show high accuracy compared to other methods (Chatfield et al. 2011). It is possible with the use of a more sophisticated classifier, for example, one that uses multiple metrics to describe the image, that the classification will be less prone to error when analysing particles with few key features.

4.3. ROC curves

As shown in Figure 7, the ROC curves for large spheres, fibrils and adhesive particles are all >90%; demonstrating the SVM is highly capable, correctly identifying debris of these morphologies without erroneously including incorrect matches. The AUC for sheets and chips was >80%.

4.4. Applications to biomedical engineering

The field of biomedical engineering places great importance on wear debris analysis, since the debris have such a pronounced effect on the life of an implant (Harris 1995; Green et al. 2000; Ingham & Fisher 2000). However, comparisons between papers from different research groups are challenging – both due to the subjective nature of debris characterisation and the variety of methodologies used. Some efforts have been made to create debris quantifiers, but these have yet to be adopted by the community as a whole, despite the methods having been published for some time (Kumar et al. 2013). It is the intention that by demonstrating the viability of using freely available machine learning and computer vision techniques developed by specialists in computer science, tribologists will be able to produce and reproduce comparable results without the need to perform redundant development of complex computer algorithms to analyse images.

5. Conclusion

This paper has investigated the accuracy of using an SVM classifier to characterise SEM images quantised with SIFT descriptors.

- The overall accuracy was 77.6 ± 4.6%.
- The overall accuracy and rates of misclassification are comparable with general computer vision papers – demonstrating these methods are suitable for wear debris analysis.
- Debris classes that contained particles which had many morphological features were classified at a greater rate with fewer false positives than classes with fewer features.
- By factoring size and AR data into the SVM, a general particle classifier could easily be constructed, making comparisons between studies possible.

Disclosure statement

Funding

ORCID

Karl D. Dearn http://orcid.org/0000-0002-8664-4303

References

Anderson DP. 1982. Wear particle atlas. Revised (Tech Rep NAEC-92-163). Lakehurst (NJ): NAEC Advanced Technology Office; p. 163.

Bentley JL. 1975. Multidimensional binary search trees used for associative searching. Commun ACM. 18:509–517.

Boiman O, Shechtman E, Irani M. 2008. In defense of nearest-neighbor based image classification. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition; Jun. Anchorage: IEEE; p. 1–8.

Bosch A, Zisserman A, Mu X, Munoz X. 2007. Image classification using random forests and ferns. Rio de Janeiro: ICCV; p. 1–8.

British Standards Institution. 2011a. BS ISO 17853:2011 wear of implant materials – polymer and metal wear particles – isolation and characterization. London: British Standards Institute.

British Standards Institution. 2011b. BS ISO 18192-1:2011 implants in surgery – wear of total intervertebral spinal disc prostheses part 1: loading and displacement parameters for wear testing and corresponding environmental conditions for test.

Canny J. 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 8:679–698.

Chatfield K, Lempitsky V, Vedaldi A, Zisserman A. 2011. The devil is in the details: an evaluation of recent feature encoding methods. In: BMVC. 2: p. 8.

Chen S, Li Z, Xu Q. 2006. Grey target theory based equipment condition monitoring and wear mode recognition. Wear. 260:438–449.

Cortes C, Vapnik V. 1995. Support-vector networks. Mach Learn. 20:273–297.

Eckold DG, Dearn KD, Shepherd DET. 2015. The evolution of polymer wear debris from total disc arthroplasty. Biotribology. 1:42–50.

Everingham M, Eslami SA, Van Gool L, Williams CK, Winn J, Zisserman A. 2015. The pascal visual object classes challenge: a retrospective. Int J Comput Vision. 111:98–136.

Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A. 2010. The pascal visual object classes (VOC) challenge. Int J Comput Vision. 88:303–338.

Green TR, Fisher J, Matthews JB, Stone MH, Ingham E. 2000. Effect of size and dose on bone resorption activity of macrophages by *in vitro* clinically relevant ultra high molecular weight polyethylene particles. J Biomed Mater Res. 53:490–497.

Green TR, Fisher J, Stone M, Wroblewski BM, Ingham E. 1998. Polyethylene particles of a 'critical size'are necessary for the induction of cytokines by macrophages *in vitro*. Biomaterials. 19:2297–2302.

Hallab NJ, Jacobs JJ. 2009. Biologic effects of implant debris. Bull NYU Hosp Joint Dis. 67:182–188.

Harris WH. 1995. The problem is osteolysis. Clin Orthop Relat Res. 311:46–53.

Hartigan JA, Wong MA. 1978. Algorithm AS 136: a k-Means clustering algorithm. Appl Statist. 28:100–108.

Hongtao L, Shirong G, Shoufan C, Shibo W. 2011. Comparison of wear debris generated from ultra high molecular weight polyethylene *in vivo* and in artificial joint simulator. Wear. 271:647–652.

Ingham E, Fisher J. 2000. Biological reactions to wear debris in total joint replacement. Proc Inst Mech Eng Part H J Eng Med. 214:21–37.

Kirk T, Panzera D, Anamalay R, Xu Z. 1995. Computer image analysis of wear debris for machine condition monitoring and fault diagnosis. Wear. 181–183:717–722.

Kumar M, Mukherjee PS, Misra NM. 2013. Advancement and current status of wear debris analysis for machine condition monitoring: a review. Ind Lubr Tribol. 65:3–11.

Lee PM. 2012. Bayesian statistics: an introduction. London: Wiley.

Lowe DG. 2004. Distinctive image features from scale-invariant keypoints. Int J Comput Vision. 60:91–110.

Mikolajczyk K, Schmid, C. 2005. A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence. 27: 1615–1630.

Moghadas P, Mahomed A, Shepherd DE, Hukins DW. 2015. Wear of the charité® lumbar intervertebral disc replacement investigated using an electro-mechanical spine simulator. Proc Inst Mech Eng Part H J Eng Med. 229:264–268.

Neukamp M, Roeder C, Veruva SY, MacDonald DW, Kurtz SM, Steinbeck MJ. 2014. In vivo compatibility of Dynesys® spinal implants: a case series of five retrieved periprosthetic tissue samples and corresponding implants. Eur Spine J. 24:1074–1084.

Nine M, Choudhury D, Hee A, Mootanah R, Osman N. 2014. Wear debris characterization and corresponding biological response: artificial hip and knee joints. Materials. 7:980–1016.

Podsiadlo P, Stachowiak GW. 2000. Scale-invariant analysis of wear particle morphology – a preliminary study. Tribol Int. 33:289–295.

Punt I, Baxter R, Ooij AV, Willems P, Rhijn LV, Kurtz S, Steinbeck M. 2011. Submicron sized ultra-high molecular weight polyethylene wear particle analysis from revised SB Charite III total disc replacements. Acta Biomater. 7:3404–3411.

Purdue PE, Koulouvaris P, Nestor BJ, Sculco TP. 2006. The central role of wear debris in periprosthetic osteolysis. HSS J. 2:102–113.

Saikko V, Vuorinen V, Revitzer H. 2015. Analysis of UHMWPE wear particles produced in the simulation of hip and knee wear mechanisms with the RandomPOD system. Biotribology. 1:30–34.

Stachowiak GW, Podsiadlo P. 1999. Surface characterization of wear particles. Wear. 225–229:1171–1185.

Stachowiak GP, Stachowiak GW, Podsiadlo P. 2008. Automated classification of wear particles based on their surface texture and shape features. Tribol Int. 41:34–43.

Vedaldi A, Fulkerson B. 2010a. [cited 2015 May 10]. Available from: http://www.vlfeat.org/applications/apps.html.

Vedaldi A, Fulkerson B. 2010b. vlfeat – an open and portable library of computer vision algorithms. [cited 2015 May 10]. Available from: http://www.vlfeat.org/applications/apps.html.

Vedaldi A, Fulkerson B. 2010c. Vlfeat: an open and portable library of computer vision algorithms. In: Proceedings of the International Conference on Multimedia, MM '10; New York, NY: ACM; p. 1469–1472.

Vedaldi A, Zisserman A. 2012. Efficient additive kernels via explicit feature maps. IEEE Transactions on Pattern Analysis and Machine Intelligence. 34:480–492.

Xin H, Shepherd DET, Dearn KD. 2013. A tribological assessment of a PEEK based self-mating total cervical disc replacement. Wear. 303:473–479.

Yang SY, Ren W, Park Y, Sieving A, Hsu S, Nasser S, Wooley PH. 2002. Diverse cellular and apoptotic responses to variant shapes of UHMWPE particles in a murine model of inflammation. Biomaterials. 23:3535–3543.

Zhang H, Berg A, Maire M, Mali J. 2006. SVM-KNN: discriminative nearest neighbor classification for visual category recognition. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition – volume 2 (CVPR'06); New York, NY: IEEE; 2: p. 2126–2136.

Zhang MQ, Lu ZP, Friedrich K. 1997. On the wear debris of polyetheretherketone: fractal dimensions in relation to wear mechanisms. Tribol Int. 30:87–102.