Using Machine Learning for Automatic Estimation of *M. smegmatis* Cell Count from Fluorescence Microscopy Images

Daniel Vente[1], Ognjen Arandjelović[2,3], Vincent O. Baron[2], Evelin Dombay[2], and Stephen H. Gillespie[2]

[1] Cardiff University, Cardiff CF10 3AT, Wales, United Kingdom
danvente@gmail.com
[2] University of St Andrews, St Andrews KY16 9SX, Scotland, United Kingdom
[3] ognjen.arandjelovic@gmail.com
https://oa7.host.cs.st-andrews.ac.uk/

Abstract. Relapse in Tuberculosis (TB) patients represents an important challenge to improve treatment. A large number of patients undergo relapse even after what was thought to be a successful treatment. Lipid rich (LR) bacteria, surviving treatment, are thought to play a key role in patient relapse. The presence of bacteria with intracellular lipid bodies in patients sputum was linked to higher risk of poor treatment outcome. LR bacteria can be stained and detected using fluorescence microscopy. However, manual counting of bacteria makes this method too labour intensive and potentially biased to be routinely used in practice or to foster large-scale data sets which would inform and drive future research efforts. In this paper we propose a new algorithm for automatic estimation of the number of bacteria present in images generated with fluorescence microscopy. Our approach comprises elements of image processing, computer vision and machine learning. We demonstrated the effectiveness of the method by testing it on fluorescence microscopy images of *in vitro* grown *M. smegmatis* cells stained with Nile red.

Keywords: Microscopy · Tuberculosis · Computer Vision · Health care · Public health · Medicine · AI.

1 Introduction

Tuberculosis (TB), a chronic pulmonary infection caused by the organism *Mycobacterium tuberculosis* (*Mtb*), is the most important cause of preventable infectious disease death. Worldwide, TB kills an estimated 1 million people annually. The majority of the impact of the disease is felt in low and middle income countries, especially in southern Africa and south-east Asia. While the WHO has resolved to end TB by 2030, relatively little progress has been made in the past decade [30].

2 Background and context

In this section, we present the relevant medical background needed to understand the motivation behind our work and the context of the broader problem thus addressed.

2.1 Lethality of *Mtb*

Mtb is an airborne bacterium which requires large amounts of oxygen to survive and is therefore predominantly found in the lungs of humans and occasionally other mammals. Due to the airborne nature of the bacterium, it can spread quickly in densely populated areas. One of the traits that make *Mtb* so difficult to treat is that after infecting a patient it can be dormant for years before actively causing the disease (the patients do not experience symptoms and transmit the disease at a low level) i.e. be in the state known as latent TB. Mycobacterial dormancy corresponds to a cell state in which bacteria exhibit low metabolic activity, the accumulation of intracellular lipid bodies, the inability to grow on solid media, and the loss of acid fastness among other features [21]. Dormant bacteria can then become active years after the first infection when the patient's immune system is weakened. Subpopulations with compromised immune systems, such as heavy smokers, people suffering from HIV, malnutrition, or diabetes are at greatly increased risk of showing active symptoms of TB [7]. Once patients declare active sensitive TB they undergo a standard six month long treatment using four antibiotics: rifampicin (RIF), isoniazid (INH), ethambutol (EMB), and pyrazinamide (PZA). The WHO defined the different objectives for TB treatment as: curing TB patients and restoring their productivity and quality of life, preventing death due to active TB or its late effects, reducing the transmission of TB, preventing drug resistance and the transmission of drug-resistant strains, and finally preventing relapse [29]. There is a strong need in particular to reduce the duration of treatment. However, new regimens tested in recent clinical trials, aiming at shortening treatment, have failed to show superiority compared to the current practice, mainly because of higher relapse rates [12, 16, 22].

2.2 Research relevance

Relapse in TB could be defined as a patient with recurrent TB symptoms becoming culture positive after having been culture negative and after completing an anti-TB treatment. In addition the original and the new isolates must have matching genotypes to confirm relapse and exclude re-infection [15]. TB relapse remains relatively poorly understood; it has been shown that relapse could occur even in patients who cleared their sputum early in treatment [24]. Bacteria showing intracellular storage of non-polar lipids represent a phenotype called lipid-rich (LR) cells, as opposed to lipid-poor (LP) cells. It is believed that LR bacteria that survive treatment play a key role in patient relapse [24]. LR cells have been shown to be up to 40 times more resistant to first-line drugs than

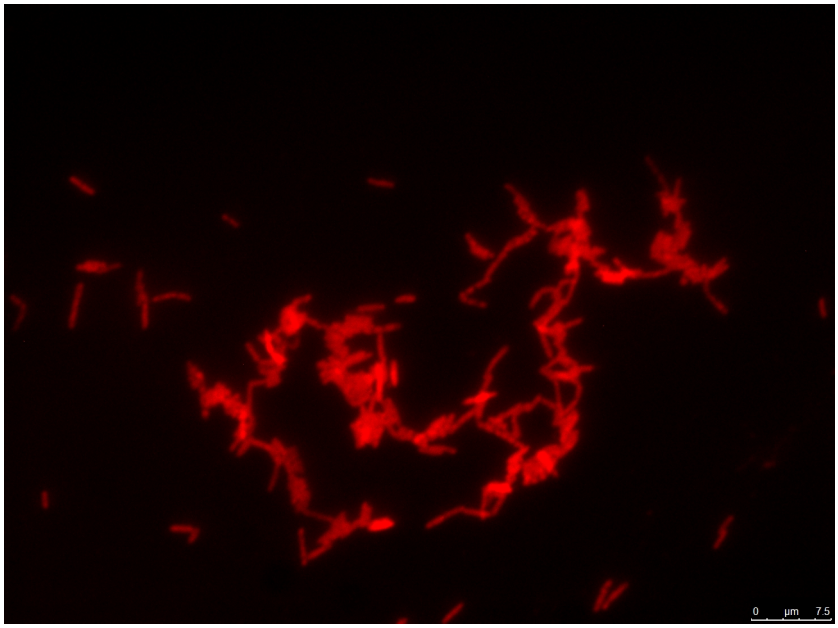


Fig. 1. Typical fluorescence microscopy image showing polar lipids of 7 day old Nile red stained *M. smegmatis* cells.

LP bacteria [14] and the presence of cells with intracellular lipid bodies in patients sputum on days 21 and 28 of treatment is associated with higher risk of poor treatment outcome [27]. This line of inquiry is highly relevant for both researchers and clinicians, as being able to detect the different bacterial phenotypes could potentially help identify patients that are at a higher risk of poor treatment outcome so they can be more carefully monitored and treated. Both polar and non-polar lipids can be detected by Nile red staining. The fluorescent properties of the fluorophore change based on whether it is located in a relatively polar or non-polar lipid environment [13]. In samples stained with Nile red, short excitation and emission wavelengths favour the detection of non-polar lipids such as triacylglycerols while higher excitation and emission wavelengths allow the detection of polar lipids (phospholipids of the membrane for example) [25]. A fluorescence microscopy image showing the polar lipids of Nile red stained *M. smegmatis* cells is shown in Fig 1. Intracellular non-polar lipid bodies, in mycobacteria, can be detected using Nile red staining and fluorescence microscopy [11]. Several previous studies have used Nile red staining to investigate the presence or absence of non-polar lipids in mycobacteria [3, 19, 14, 8, 9, 18]. Manual counting of bacteria is a very labour intensive process. The first step in developing a software solution to report on the relative percentages of LR and LP cells present is to count the total number of cells. Therefore this paper proposes an automatic procedure for estimating the cell number present in fluorescence images of *M. smegmatis* cells stained with Nile red.

3 Technical details

As can be seen in Figs 1 and 2, the difficulty in counting the bacteria emerges from the fact that they are often densely packed or even overlapping, so that it can be difficult to distinguish them individually. Thus, to summarize, our method approaches the task in several steps to address different challenges. First, we employ Canny edge detection (CED) and morphological image processing to identify key image areas of interest (AOI). After AOI are identified, features based on Local Binary Patterns are extracted and used to describe the corresponding content. Finally, rather than attempting to count individual cells, a regression based approach is used for the inference of the cell count in each AOI. An overview of the process can be seen in Fig 3.

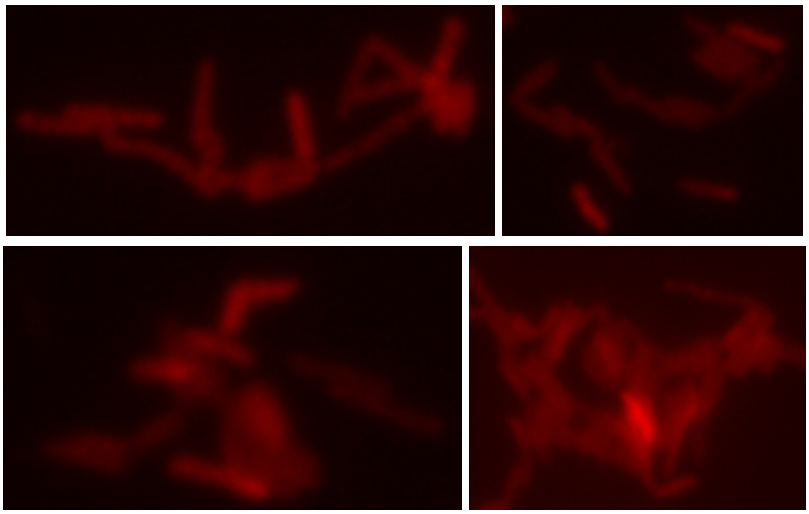


Fig. 2. Magnified input image patches exemplifying the impracticability of counting individual cells.

3.1 Data acquisition

Bacterial culture *M. smegmatis* (NCTC 8159), was grown at 37°C in Middlebrook 7H9 medium (Sigma-Aldrich), supplemented with 0.45% (v/v) of glycerol (Sigma-Aldrich) and with 0.05% (v/v) Tween80 (Fisher Scientific).

Sample preparation In this work two experiments were performed. Both comprise two sets of prepared samples: an early exponential phase culture (24-hour-old) and a stationary phase culture (7-day-old). In each experiment, two times 100 μl from a 7-day-old culture were taken and stained with Nile red. At the same time 200 μl of the 7-day-old culture was spun down (20,000 g for 3 minutes) and

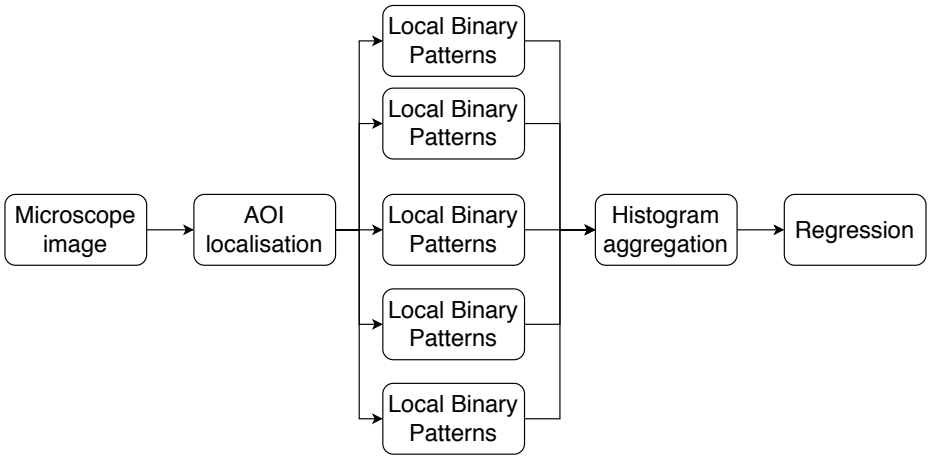


Fig. 3. High level conceptual overview of the proposed algorithm.

then resuspended in 500 μl of fresh 7H9 medium. The bacterial suspension was incubated at 37°C for 24 hours. Then the culture tube was spun down (20,000 g for 3 minutes), the supernatant removed and the pellet resuspended in 200 μl of phosphate buffered saline (PBS). Two times 90 μl from this suspension were taken and stained with Nile red.

Nile red staining Using a Nile red (Sigma-Aldrich) stock solution at 250 $\mu g/ml$ dissolved in dimethyl sulfoxide (DMSO), 0.9 to 1 μl was added to the bacterial suspension (90 μl or 100 μl) to obtain a final Nile red concentration of 2.5 $\mu g/ml$. The tubes were then vortexed and left in the dark (covered with aluminium foil) at room temperature for 10 minutes. The bacterial suspensions were then centrifuged at 20,000 g for 3 minutes and the supernatant was discarded. Following this the bacteria were washed twice using PBS (the pellet was resuspended in PBS, the tubes vortexed, then the tubes were centrifuged at 20,000 g for 3 minutes and the supernatant was discarded). Finally, the bacterial pellets were resuspended in 20 μl of PBS and 10 μl was heat fixed on top of a microscopy slide. Similar Nile red staining protocols have been used with success in previously published work [3, 19].

Fluorescence micropscopy The microscopy slides were then observed using a fluorescence microscope (Leica DM5500). The objective used was a 100×-magnification oil immersion objective. A Leica camera DFC 3000 G was used to capture images. A L5 filter cube, presenting an excitation of 480/40 nm and an emission of 527/30 nm was used to observe the fluorescence from Nile red located in a non-polar lipid environment. The TX2 filter cube with an excitation light of 560/40 nm and an emission of 645/75 nm was used to detect the fluorescence from Nile red present in a polar lipid environment.

3.2 Proposed method

Localization The first step in the process is to obtain an AOI on which to perform our learning. An example of this process is shown in Figure 5. Firstly, the image was preprocessed using contrast stretching, a form of intensity normalization that is applied as follows:

$$I_{out} = (I_{in} - p_{low})\frac{255}{p_{high} - p_{low}} \tag{1}$$

where p_{low} and p_{high} represent lower and higher percentiles which are preset algorithm parameters.

After that, Canny edge detection [6] is used to produce a binary image which captures variable information content across the input image [2]. To summarize the key ideas, CED applies a Gaussian blur to the image to reduce high frequency noise. Then the Sobel operator [26] is applied as a means of approximating the intensity gradient at each pixel.

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

Fig. 4. Sobel edge detection directional kernels, G_y and G_x respectively.

The Sobel operator comprises the application of two kernels, G_x and G_y respectively as can be seen in Fig 4. As per the convolution theorem:

$$I * G = \mathcal{F}^{-1}\{\mathcal{F}\{I\} \cdot \mathcal{F}\{G\}\} \tag{2}$$

where $\mathcal{F}$ denotes the Fourier transform [26]. We can then calculate the magnitude and orientation of the gradient as follows:

$$\mathbf{G} = \sqrt{(G_x * I)^2 + (G_y * I)^2} \tag{3}$$

$$\Theta = \arctan((G_y * I)/(G_x * I)). \tag{4}$$

Non-maximum suppression to the gradient magnitude is applied thereafter. Since the intensity gradient was calculated in the previous step and we are looking for a continuous edge, only the pixels in the direction of the gradient and the negative gradient have to be checked. Non-local maxima pixels are set to 0. Finally, hysteresis thresholding is applied. This means that two thresholds are made use of. Any pixel with an intensity below the lower threshold is set to

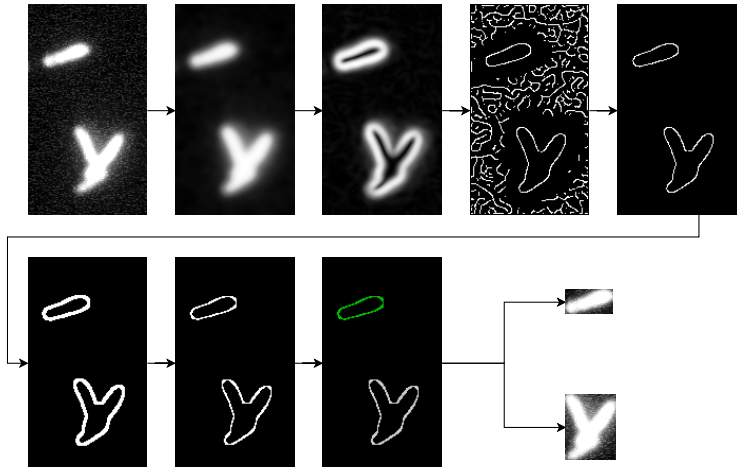


Fig. 5. Illustrative example of the proposed AOI extraction. One of the detected regions contains a single bacterial cell whereas the other contains multiple cells which are not readily distinguished one from another without expert semantic knowledge.

0, and anything that has a gradient above the upper threshold is set to 1. If a gradient falls between the thresholds, it is set to 1 only if it neighbours an edge pixel.

For the AOI extraction step, it is important that a bacterial clump is detected in its entirety, as one connected object. Due to possibly non-uniform lighting, focusing problems, as well as other potential issues encountered during image acquisition, it is possible that the edge detector produces breaks in salient edges. For this reason, repeated morphological dilation is applied to the original binary images, thickening edges and thus closing small edge breaks. However, this introduces a tradeoff: if an edge is dilated too much it can merge with neighbouring clusters. To minimize this effect, erosion is applied after each dilation, thus thinning the edge again, while retaining its greater continuity. Note that combining dilation and erosion with 8- and 4-connectivity respectively has a smoothing effect. After these operations, connected component labelling is applied to the produced binary image. Then we extract the coordinates of the bounding box from the binary image, which we then use to crop out the AOI from the original images [26].

Feature extraction After AOI are localized, each is represented by a histogram of local binary patterns (LBPs) [17, 10]. A local binary pattern is parameterized by two values, P and R respectively, which represent the number of points sampled, and the distance at which they are sampled from the target locus pixel.

The corresponding local binary pattern is then:

$$LBP_{P,R}(x) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \tag{5}$$

with:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where g_c is the intensity of the centre pixel and g_p is the intensity of the p-th pixel at distance R. A single LBP is readily represented by a number, as illustrated in Fig 6, and an image patch by the corresponding histogram, as shown in Fig 7.

The eventual feature vector used in the regression step is the sum of all the histograms of the local binary patterns that are produced by processing all the AOI in a sample. This feature vector is then used by a regression algorithm to produce a prediction [1].

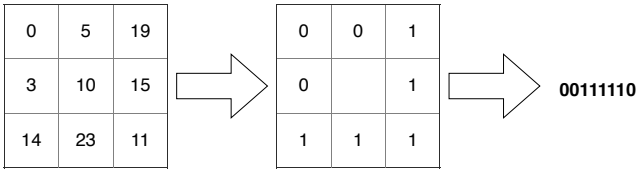


Fig. 6. Example of $LBP_{8,1}$ extraction for an elementary image patch.

3.3 Model selection and parameter learning

Since the number of cells in a particular slide can vary significantly, we decided to use relative metrics. In particular we evaluated the models using the Percentage Error (PE) and the Mean Absolute Percentage Error (MAPE).

Following the successes of such approaches reported in the recent literature [23], we considered several regression types, namely linear (LR) [20], neural network based (NN) [4], decision tree based (DT) [31], gradient boosting machine based (GB) [28], and random forest based (RF) [5], the best amongst them selected automatically. We imposed appropriate distributions over the parameters of all the algorithms [23], and let each configuration run a randomized parameter search of 1000 iterations, using 3-fold cross-validation for statistical robustness. The inferred best model was used for the final error analysis.

4 Results and discussion

The results of the randomized parameter search are summarized in Table 1. In short, the gradient boosting based approach significantly outperforms the alternatives included in the selection process. Therefore this regression methodology was adopted for use in the final analysis presented hereafter.

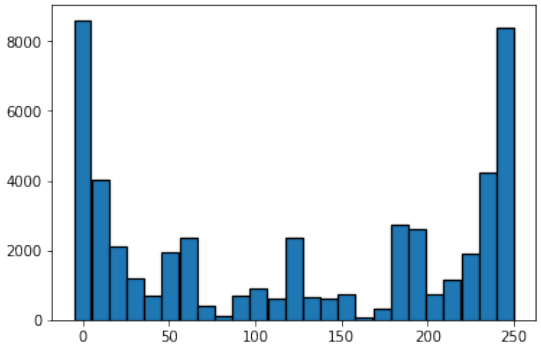
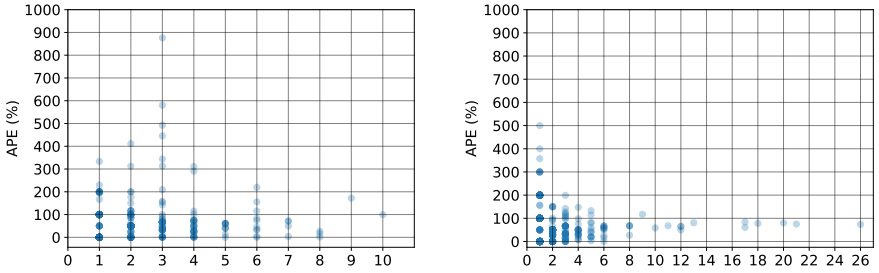


Fig. 7. LBP histogram example ($P = 8$, $R = 1$).

Table 1. Random parameter search results.

	LR	NN	DT	GB	RF
MAPE	0.347	0.322	0.249	**0.055**	0.242

To gain insight into the overall structure of the proposed method's performance we started our analysis by examining the dependence of the error on the true, target number of cells within a specific area of interest. The corresponding plots for the two experiments are shown in Figs 8(a) and 8(b). It can be seen that most of the overall error is contributed to by a small number of areas of interest. It is even more important to observe that these generally correspond to areas with a small cell count – considering that we are looking at relative rather than absolute error, this is reassuring because it suggests low overall absolute error (error for the entire input image or slide).



(a) Experiment 1 (b) Experiment 2

Fig. 8. Prediction error as a function of the true cell count.

To examine our hypothesis, we next looked at the slide level errors – the corresponding results are summarized in Table 2. As the figures clearly show,

our method's performance is outstanding, resulting in the slide level error of less than 6.5%. Interpreted together with the previously discussed results, these statistics demonstrate both the relative insignificance of the somewhat higher proportional error rate for sparsely populated areas of interest and, importantly, that the errors seem to be symmetrically distributed, leading to cancellation of overcounts and undercounts when aggregated over an entire input image.

Table 2. Summary of experimental results.

	Predicted count	True count	Difference	Relative error (%)
Experiment 1	986	1053	67	6.3
Experiment 2	1015	1020	5	0.4

5 Summary and conclusions

TB remains a global health issue worldwide and relapse in TB patients is a major obstacle to improving treatment conditions. LR cells are believed to play a central role in relapse. The presence of cells with intracellular lipid bodies in patients sputum was associated with higher risk of poor treatment outcome. Therefore, the proportion of LR cells in patients sputum samples in early treatment could be an indicator of long term treatment outcome.

In this paper we proposed an automatic method for estimating the number of the bacteria present in a fluorescence microscopy image. Our method uses Canny edge detection, morphological image processing, and connected component labelling to extract salient image regions, the content of which is then captured by local binary pattern histograms, followed by a machine learning stage which learns the mapping from interest region representations to cell counts. Using data sets, generated from *in vitro M. smegmatis* cultures, we demonstrated that the proposed model performs extremely well, achieving less than 6.5% error. These results provide strong evidence of the potential of automatic image analysis tools for stained sputum smears and motivate further work in the area.

Our immediate follow-up work will focus on extending the method to the estimation of LR cell count. In addition, we intend to extend the method so that it can deal with patient samples which demand the ability to distinguish between bacteria and confounding material found in this type of data.

References

1. Arandjelović, O.: Reimagining the central challenge of face recognition: turning a problem into an advantage. Pattern Recognition pp. 388–400 (2018)
2. Arandjelović, O., Cipolla, R.: A new look at filtering techniques for illumination invariance in automatic face recognition. In Proc. IEEE International Conference on Automatic Face and Gesture Recognition pp. 449–454 (2006)

3. Baron, V.O., Chen, M., Clark, S.O., Williams, A., Hammond, R.J., Dholakia, K., Gillespie, S.H.: Label-free optical vibrational spectroscopy to detect the metabolic state of M. tuberculosis cells at the site of disease. Scientific Reports **7**(1), 1–9 (2017)
4. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, Oxford, England (1995)
5. Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001)
6. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence **8**(6), 679–698 (1986)
7. Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., S.Gas, III, C.E.B., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G.: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature **396**(6685), 1–27 (1998)
8. Daniel, J., Kapoor, N., Sirakova, T., Sinha, R., Kolattukudy, P.: The perilipin-like PPE15 protein in Mycobacterium tuberculosis is required for triacylglycerol accumulation under dormancy-inducing conditions. Molecular Microbiology **101**(5), 784–794 (2016)
9. Daniel, J., Maamar, H., Deb, C., Sirakova, T.D., Kolattukudy, P.E.: Mycobacterium tuberculosis uses host triacylglycerol to accumulate lipid droplets and acquires a dormancy-like phenotype in lipid-loaded macrophages. PLoS Pathogens **7**(6) (2011)
10. Fan, J., Arandjelović, O.: Employing domain specific discriminative information to address inherent limitations of the LBP descriptor in face recognition. In Proc. IEEE International Joint Conference on Neural Networks (2018)
11. Garton, N.J., Christensen, H., Minnikin, D.E., Adegbola, R.A., Barer, M.R.: Intracellular lipophilic inclusions of mycobacteria in vitro and in sputum. Microbiology **148**(10), 2951–2958 (2002)
12. Gillespie, S.H., Crook, A.M., McHugh, T.D., Mendel, C.M., Meredith, S.K., Murray, S.R., Pappas, F., Phillips, P.P.J., Nunn, A.J.: Four-month moxifloxacin-based regimens for drug-sensitive tuberculosis. New England Journal of Medicine **371**(17), 1577–1587 (2014)
13. Greenspan, P., Fowler, S.D.: Spectrofluorometric studies of the lipid probe, Nile Red . Journal of Lipid Research **26**(7), 781–789 (1985)
14. Hammond, R.J., Baron, V.O., Oravcova, K., Lipworth, S., Gillespie, S.H.: Phenotypic resistance in mycobacteria: is it because I am old or fat that I resist you? Journal of Antimicrobial Chemotherapy **70**(10), 2823–2827 (2015)
15. Jasmer, R.M., Bozeman, L., Schwartzman, K., Cave, M.D., Saukkonen, J.J., Metchock, B., Khan, A., Burman, W.J.: Recurrent tuberculosis in the United States and Canada: relapse or reinfection? American Journal of Respiratory and Critical Care Medicine **170**(12), 1360–1366 (2004)
16. Jindani, A., Harrison, T.S., Nunn, A.J., Phillips, P.P.J., Churchyard, G.J., Charalambous, S., Hatherill, M., Geldenhuys, H., McIlleron, H.M., Zvada, S.P., Mungofa, S., Shah, N.A., Zizhou, S., Magweta, L., Shepherd, J., Nyirenda, S., van Dijk, J.H., Clouting, H.E., Coleman, D., Bateson, A.L.E., McHugh, T.D., Butcher, P.D., Mitchison, D.A.: High-dose rifapentine with moxifloxacin for pulmonary tuberculosis. New England Journal of Medicine **371**(17), 1599–1608 (2014)

17. Karsten, J., Arandjelović, O.: Automatic vertebrae localization from CT scans using volumetric descriptors. In Proc. International Conference of the IEEE Engineering in Medicine and Biology Society pp. 576–579 (2017)
18. Kayigire, X.A., Friedrich, S.O., Van Der Merwe, L., Donald, P.R., Diacon, A.H.: Simultaneous staining of sputum smears for acid-fast and lipid-containing Myobacterium tuberculosis can enhance the clinical evaluation of antituberculosis treatments. Tuberculosis **95**(6), 770–779 (2015)
19. Kennedy, J.A., Baron, V.O., Hammond, R.J., Sloan, D.J., Gillespie, S.H.: Centrifugation and decontamination procedures selectively impair recovery of important populations in Mycobacterium smegmatis. Tuberculosis **112**, 79–82 (2018)
20. Li, J., Arandjelović, O.: Glycaemic index prediction: a pilot study of data linkage challenges and the application of machine learning. In Proc. IEEE International Conference on Biomedical and Health Informatics pp. 357–360 (2017)
21. Lipworth, S., Hammond, R.J., Baron, V.O., Hu, Y., Coates, A., Gillespie, S.H.: Defining dormancy in mycobacterial disease. Tuberculosis **99**, 131–142 (2016)
22. Merle, C.S., Fielding, K., Sow, O.B., Gninafon, M., Lo, M.B., Mthiyane, T., Odhiambo, J., Amukoye, E., Bah, B., Kassa, F., N'Diaye, A., Rustomjee, R., de Jong, B.C., Horton, J., Perronne, C., Sismanidis, C., Lapujade, O., Olliaro, P.L., Lienhardt, C.: A four-month gatifloxacin-containing regimen for treating tuberculosis. New England Journal of Medicine **371**(17), 1588–1598 (2014)
23. Neofytos, D., Arandjelović, O., Harrison, D., Caie, P.D.: Machine learning based prognosis of stage II colorectal cancer outcome. npj Digital Medicine (2018)
24. Phillips, P.P., Mendel, C.M., Burger, D.A., Crook, A., Nunn, A.J., Dawson, R., Diacon, A.H., Gillespie, S.H.: Limited role of culture conversion for decision-making in individual patient care and for advancing novel regimens to confirmatory clinical trials. BMC Medicine **14**(1), 1–11 (2016)
25. Rumin, J., Bonnefond, H., Saint-Jean, B., Rouxel, C., Sciandra, A., Bernard, O., Cadoret, J.P., Bougaran, G.: The use of fluorescent Nile red and BODIPY for lipid measurement in microalgae. Biotechnology for Biofuels **8**(1), 1–16 (2015)
26. Shapiro, L., Stockman, G.: Computer Vision. Pearson (2000)
27. Sloan, D.J., Mwandumba, H.C., Garton, N.J., Khoo, S.H., Butterworth, A.E., Allain, T.J., Heyderman, R.S., Corbett, E.L., Barer, M.R., Davies, G.R.: Pharmacodynamic modeling of bacillary elimination rates and detection of bacterial lipid bodies in sputum to predict and understand outcomes in treatment of pulmonary tuberculosis. Clinical Infectious Diseases **61**(1), 1–8 (2015)
28. Tun, W., Arandjelović, O., Caie, D.P.: Using machine learning and urine cytology for bladder cancer prescreening and patient stratification. In Proc. AAAI Conference on Artificial Intelligence Workshop on Health Intelligence pp. 507–513 (2018)
29. World Health Organization: The Treatment of Tuberculosis: Guidelines. World Health Organization, Geneva (2010)
30. World Health Organization: WHO — Top 10 causes of death (2018)
31. Zadrozny, B., Elkan, C.: Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. In Proc. IMLS International Conference on Machine Learning **1**, 609–616 (2001)