Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 192 (2021) 2002–2011

Procedia Computer Science

www.elsevier.com/locate/procedia

25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

A Deep Learning-based Surrogate for the XRF Approximation of Elemental Composition within Archaeological Artefacts before Restoration

Ruxandra Stoean[a,b], Leonard Ionescu[a,c], Catalin Stoean[a,b,**], Marinela Boicea[c], Miguel Atencia[d], Gonzalo Joya[d]

[a]*Romanian Institute of Science and Technology, Cluj-Napoca, Romania*
[b]*Faculty of Sciences, University of Craiova, Romania*
[c]*Restoration and Conservation Lab, Oltenia Museum, History and Archaeology Section, Craiova, Romania*
[d]*Universidad de Málaga, Málaga, Spain*

Abstract

The restoration of archaeological artefacts is naturally utterly important for preserving the cultural heritage. The first step that is undertaken in this process is the chemical analysis of the object, in order to decide the best procedures for its restoration. The gold standard in approximating the concentration of the elements in its composition (in percentages, between 0 and 100) is performed through an X-ray fluorescence (XRF) machine. While this is a non-invasive approach, it comes at substantial financial and training costs, and possible radiation exposure of the investigator. In this context, the present paper explores the potential of a deep learning regression model to give an estimate on the concentration of a given element from stereo microscopy slides of historical artefacts, as an alternative means to the XRF. Two problems with different degrees of complexity are examined in turn. The first one is represented by the consideration of iron objects, where the metal is strongly dominant in the chemical structure. The second comes both as a complement to the other, in order to expose the model also to non-iron items, and as a more difficult task of identifying the degree of copper that is present only as part of an alloy constitution. While for iron the one absolute value prediction of the model is always very close to the XRF approximation, copper has a wider distribution of its concentration among objects, which is more challenging to learn; hence, performance for a singular absolute estimation can rise only with the increase in the amount of data. A window of error acceptability was also implemented and it allows for an approximation that is sufficient for grasping the degree of the metal in the composition that is necessary for the restoration procedures. The findings therefore provide a first step in putting forward a computational support tool that represents a less expensive and less dangerous alternative for approximating the elemental analysis before artefact reinstatement.

Keywords: deep learning; regression; chemical analysis; X-ray fluorescence; archaeology; restoration

* catalin.stoean@inf.ucv.ro
** Corresponding author.

1877-0509
10.1016/j.procs.2021.08.206

1. Introduction

The deterioration of the material heritage of our culture is inevitable. The assets that form our cultural heritage cannot be considered constant from a chemical point of view. Their decay thus follows in the form of various degradation processes, determined by time and the interaction between the artefacts and their environment.

Therefore, before restoring an archaeological object, a multi-faceted assessment of the composition and the degradation of its material is essential in deciding the most appropriate method of reestablishment and conservation [1]. Different (micro)invasive or partially destructive physico-chemical analysis techniques are performed to assess alloy composition, proteic nature, acidity, solubility, ink and color types. Alternatively, several non-invasive procedures for the object surface, such as optical analysis or photo-fixing can investigate corrosion, erosion, textures and watermarks. Finally, more recent and effective direct analysis approaches are reflectography, radiography, X-ray fluorescence (XRF) and trace analysis.

The XRF technology targets the chemical composition of a heritage object at the surface level and thus gives an approximation of the concentrations of the main elements that is useful for the decision regarding the restoration procedure [2]. Nevertheless, the fact that the investment in this machine is rather high and that the operators need special training, with the extra (even if only reduced) possibility of a radioactive contamination, limit its adoption as a largely used method.

Hence, the use of machine learning technologies for the examination of microscopy slides as an alternative to XRF scanning can assist the restorer in approximating the elemental composition of the artefact in a non-invasive, radiation-free, labor-effective and cost-efficient manner. A deep learning (DL) regression model acting as a surrogate for the XRF in estimating the chemical concentration of a given element from archaeological objects is therefore put forward in the current paper.

Microscopy images of iron (Fe) historical artefacts labelled with the XRF output are gathered and analyzed for detecting the metal concentration (between 0 and 100, in percents). In order to also show examples with close to 0 percents in Fe to the machine, bronze objects were also added to the collection. Bronze is an alloy of copper (Cu) and other metals, as it had been discovered in the earlier Bronze Age that the metal objects were made superior through such a combination. Additionally, the second set of samples gives rise to the option of testing the generalization ability of the method for the quantification of another metal. Moreover, this next task provides a contrasting, more challenging setting for the DL regression: since Cu is part of an alloy within the bronze pieces, its concentration has a considerably wider spread numerically within the artefacts considered, in opposition to Fe, which is predominant in its items.

The remainder of the work is structured as follows. The task of approximating the chemical concentrations of a metal on the surface of ancient objects and the two problems considered for distinct metal properties are enunciated in section 2. The state of the art in using computational support for XRF spectrometry is outlined in subsection 3.1 of the methodology, while the proposed DL model for estimating the percentage of elemental composition from microscopy images is presented in subsection 3.2. The experiments carried for each of the two problems are described in separate turns in section 4. The inclusion of a window of error acceptability that is in line with the elemental approximation needed for restoration decisions is outlined in subsection 4.3. The limitations and the prospects of the current version of the computational XRF surrogate are discussed in section 5.

2. The regression instances and the data set

All data is provided courtesy of the Oltenia Museum - the History and Archaeology Section in Craiova, Romania. All the analyzed artefacts having arrived during the months of January and February 2021 for restoration at the museum have been taken into consideration. A number of 31 iron historical items have been collected. As a complement, another number of 23 bronze objects have been gathered.

Each object is photographed under the microscope in several areas surrounding the XRF point of analysis. There are between 2 and 10 images processed for one object. A total of 398 microscopy images (251 Fe and 147 Cu) had been consequently acquired from these objects through an Olympus SZX-7 optical stereo microscope equipped with a dedicated digital camera, model Quick Photo Micro 2.2. There is no pre-processing step taken to standardize light and noise across images. The chemical analysis of the samples had been performed through a portable Bruker Titan

Fig. 1. Sample of iron artefacts with one associated microscopy image and XRF-derived chemical composition. The concentration of Fe is highlighted in bold.

Fig. 2. Sample of bronze artefacts with one associated microscopy image and XRF-derived chemical composition in terms Cu. The concentration of Cu is highlighted in bold.

S1 XRF spectrometer, used in conjunction with the dedicated Artax software. Some examples of iron artefacts, each with one of the related microscopy scans produced and composition label as found by the XRF, are given in Figure 1. The concentration is highlighted in bold. It can be seen that Fe is strongly prevailing in its related items.

Figure 2 shows bronze objects, each with one connected microscopy image and XRF output. As it can be noticed, Cu is always part of an alloy in the bronze artefacts, with a wide range of concentrations, which makes it truly intricate to assess quantitatively once the amount of available data is limited.

Since the images of one metal also serve as counterexamples in the training of the other (for the considered cases of Fe and Cu), its minor concentrations (standing for negative) are also outlined besides each image. Looking at the pictures from both figures, it can be seen that it is difficult even to visually distinguish one chosen metal from the other qualitatively, making significantly harder to grasp the difference between two (even) numerically distant absolute percentage values of the same metal in the composition without the XRF support.

3. Methodology

The present paper appoints a DL regression architecture for a fast chemical approximation of the percent of a given chemical element in archaeological artefacts from stereo microscopical slides, as an effective surrogate for the

expensive and radioactive XRF. Contrary to the related literature entries that operate only in tandem with the XRF, taking from its numerical outcomes or spectra, the proposed DL approach will look at microscopic images, taking into account the chemical concentrations estimated by the XRF only as a ground truth for training and validation purposes. The elemental concentration derived from the regression model will be confronted with the XRF output and assessed through the resulting difference.

3.1. State of the art

The current computational approaches for supporting this chemical analysis are rather scarce and generally deal with sampling or analyzing XRF spectral data. The study [3] puts forward a Monte Carlo approach to simulate the XRF spectra emitted in the case of metal artefacts. In [4], spectra simulated again by a Monte Carlo method are given to a neural network for multi-output regression, in order to quantify the elemental concentrations for five elements (Au, Ag, Cu, Ni, Zn) in a sample.

In the opposite direction, taking the chemical composition of each artefact derived from Energy Dispersive X-ray Fluorescence as independent attributes, the paper [5] appoints k-nearest neighbors, a C4.5 decision tree and learning vector quantization to classify ceramic objects into several classes.

The current work seeks to corroborate digital stereo microscopy images with the XRF composition approximation as the ground truth. The combined data is used to train a DL model for regression to learn the correspondence between the characteristics of the image and the XRF output regarding the concentration of an element. The XRF has been previously combined with micro-analyses of samples by investigators in physics and chemistry in order to gain insight into the underlying structure of art objects [6].

3.2. Proposed Approach

DL was thought to conduct image analysis through convolutional neural networks (CNN). Image inspection is typically performed towards a classification goal. However, starting with the Udacity challenge [7], where the aim was to predict the correct steering angle from forward facing images from a car, CNN have achieved interesting results for regression as well. Some recent examples include crowd counting from camera surveillance [8], ground sampling distance estimation from remote sensing devices [9] and leaf counting [10].

The regression task is nevertheless far more difficult in the current scenario, as the output is not something that can be easily quantified numerically as the problems of counting objects or measuring distances in the found literature. The percent of a metal cannot be assessed just by looking at the microscopic images, at least not by the human experts. In this sense, the scenario seems to be more similar to people age prediction from face images [11, 12]. Even as this comparison goes, people may have similar traits as they age, e.g. wrinkles, creases and specific lines. For the problem at hand, the objects are visually examined for composition, not appearance, so the quantification is still different. In addition, corrosion due to metal aging and soil conditions may also change composition detection in otherwise chemically similar objects. Moreover, the amount of face images available for training is far more easily procurable than the number of samples limited by the XRF very specific operation.

Consequently, for the current task, the convolutional layers will extract the most relevant features of the microscopic images and these will be further used in learning their correspondence with the XRF elemental quantification, having a final fully connected layer with one unit and loss being given by the mean squared error (MSE). A special attention must be given such that the resulting amount always remains in the interval $[0, 100]$, as the elemental composition output is given in percents. Images where the studied metal is present as primary are taken together with samples of other metals (where it is minor or absent), such as to train the approach on both positive and negative examples.

4. Experiments

A first regression problem is formulated for the estimation of the percentage of Fe from related artefacts. This primary experiment aims to test the viability of the DL to mimic the XRF approximation. Samples of non-Fe composition are also considered for training the network also to recognize the lack of the metal.

Table 1. Performance metrics: there are N samples, $\hat{y}_i$ is the model prediction for sample i, y_i is its XRF value and Med stands for the median

Metric	Expression
Mean squared error	$MSE(y, \hat{y}) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$
Mean absolute error	$MAE(y, \hat{y}) = \frac{1}{N}\sum_{i=1}^{N}\|y_i - \hat{y}_i\|$
Median absolute error	$MedAE(y, \hat{y}) = Med(\|y_1 - \hat{y}_1\|, ..., \|y_N - \hat{y}_N\|)$

A second regression instance is constructed for Cu objects, where the variability of the metal concentration in connected objects is very high. This second experiment targets the generalization ability of the approach as well as the data conditions needed to be met for a good learning and approximation in different scenarios.

4.1. Experiment 1: Approximation of the concentration of Fe

There are 54 pieces in total (31 iron and 23 bronze) in the data set and for each item there are 2 up to 11 microscopical images taken. The total number of images is 398 (251 Fe and 147 Cu). There are no objects that have pictures appearing in both the training/validation couple and in the test set. The training and validation set comprises pictures from 43 objects and the test set contains images from the remaining 11 items. Switching to images now, the training and validation collection contains 316 images and the test has 78. The sets contain both Fe and non-Fe samples. k-fold cross validation with $k = 5$ is performed on the 316 images and each of the 5 obtained models is subsequently applied on the test set.

A ResNet-34 architecture is chosen for the current regression task. This implies a conv(7×7 size, 64 filters), maxpool(3×3, stride 2), a first block of 3 repetitions of the sequence conv(3×3, 64), conv(3×3, 64), a second with 4 iterations of the series conv(3×3, 128), conv(3×3, 128), a third one with 6 cycles of the flow conv(3×3, 256), conv(3×3, 256) and a fourth with 3 iterations of the order conv(3×3, 512), conv(3×3, 512). The convolutional part ends with average pooling and a fully connected layer with 1000 units. Another dense layer with 1 unit (for regression) ends the structure. The network proves to be a good compromise between performance and its size, especially since the data set is relatively small and the chances of reaching overfitting are thus reduced. Other architectures have been tried in pre-experimental planning, but the results were not significantly different. A batch size of 32 is used. Data augmentation is performed again in order to reduce the possible overfitting due to the small sample size. Various options were tried for this step and the final selection was to use a random image flip, a random rotation between -180 and 180 degrees, symmetric warp of magnitude between -0.3 and 0.3, zoom of up to 1.1 times. The probability of applying each of these transformations was of 0.5. The images were resized to 224×224.

A discriminative learning rate is used, meaning that the very first layers are trained at 1e-05, the last at 1e-01, and the learning rates of the group layers in between are linearly spaced between the first, the middle (1e-04), and the last one. The model is run for 100 epochs, but having an additional stop condition at 30 epochs without improvement. A 1cycle policy [13] is used for the training process. After the initial training, all layers are unfrozen and training with the same 1cycle policy is resumed for another 10 epochs. The learning rate in the second step is fine-tuned automatically and its values are chosen according to the strategy proposed in [13].

The loss function is given by the MSE. In each of the 5-fold runs, the best model obtained on the validation set with respect to the MSE is retained and afterwards applied on the test set. In addition to the MSE, the mean absolute error (MAE) and the median absolute error (MedAE) are also calculated. The underlying formulas are revisited in Table 1. All the results are recorded as an average over 5 repetitions.

The implementation is done in Python, using Fastai and Pytorch libraries and the code is run in Google Colab using the available GPU.

4.1.1. Results and visualization

Figure 3 shows some insights about the amount of Fe that is found in the samples from the data, both in the training/validation used for 5-fold cross-validation, as well as in test set. It can now be clearly seen that Fe is present with a concentration over 60% at one end of the spectrum and in an amount not exceeding 20% at the other end.

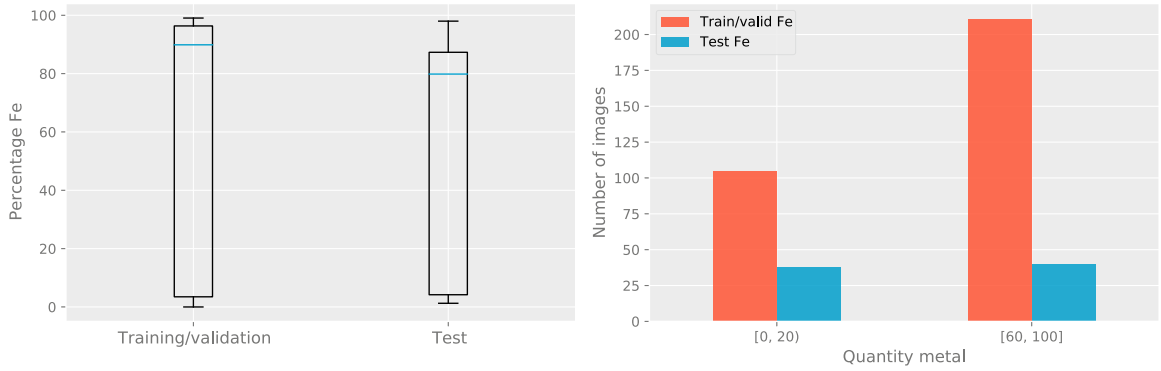

Fig. 3. The left illustration shows box plots with the percentage of Fe for the samples in the training and validation sets, as used in the 5-fold cross-validation, and in the test set. The right plot shows the number of images that contain Fe in two interval concentrations.

Table 2. Mean validation/test metrics obtained by the model after the first stage of training and the subsequent one on recognizing the Fe amount.

Set & variant	MSE	MAE	MedAE
Validation frozen	36.93	4.45	3.59
Validation unfrozen & retrained	37.61	4.65	3.59
Test frozen	82.33	6.14	4.02
Test unfrozen & retrained	70.05	6.04	4.57
Test mean over pics per object	40.49	4.85	3.86

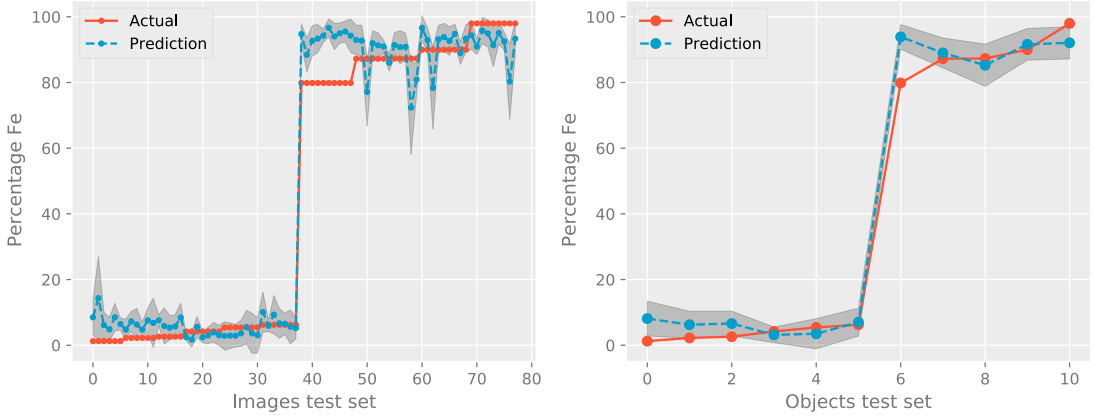

Fig. 4. XRF values for Fe (red), and average predictions (blue) with standard deviation as a shadow for the test images (left) and objects (right).

Table 2 shows the average results over the 5 runs for all the metrics discussed in the previous subsection for validation and test. The values are computed at the level of pictures, as well as for each object in part, when mean scores are calculated over all its pictures. The running time for the entire training and testing takes around 19 minutes.

Figure 4 illustrates the XRF percentage values in opposition to the predicted ones for Fe, obtained by computing the mean from the application of the model on the test set in 5 runs. The first plot contains the results for the individual pictures and the second one shows the objects with values averaged from all their images. The standard deviations for the predictions are illustrated as a grey shadow. The pictures and objects in the plots are ordered based on the XRF value. In the first plot, every dot corresponds to a picture, and, in the second plot, each point is related to an object.

4.1.2. Discussion

The left part of Figure 3 illustrates box plots with the quantity of Fe, as determined by the XRF, that appears in the images from the training/validation and test subsets. The right plot shows another mode of the disposal of samples per concentration over the subsets. The amount of samples that have Fe below 20% is shown in the first pair of bars and the second one indicates the counts for concentrations above 60%. Fe has no items where its role in the composition is between one third to half. This is natural since iron was found alone to be superior in material strength as opposed to bronze (which is an alloy), with the advancement in metallurgy that led to the rise of the Iron Age.

Table 2 illustrates the results for the Fe study case. The samples in the test set contain objects that do not have any picture in the training/validation folds. However, the 5-fold cross validation is applied directly on the images, without separating the objects, as well, into training and validation. This represents an explanation for the slightly better results for validation, as opposed to the test outputs.

The difference between the test frozen results and test unfrozen and retrained ones is that for the latter all layers are unfrozen, the optimal learning rate is searched for and the 1cycle policy is applied for another 10 epochs. This second step takes around 4 minutes out of the total of 19 for the entire process and, in general, significantly improves the results.

The results for Fe are very good for both validation and test and, although the additional training of the model leads to slightly better results, the improvement is not significant. MSE is the metric that is used for the loss of the CNN. While the unfreezing and additional training did not change the validation results, for the test samples the error output decreased from 82.33 to 70.05. From the perspective of the archaeological experts, the MAE and MedAE indicate best how close the model is in predictions to the values of the XRF. The MedAE value is more robust in this case, as it is insensitive to the outliers generated from learning after multiple microscopic facets are taken for the same surface but at closely distant spots. The additional training led in these cases did not further improve results.

The last row from Fe in Table 2 contains results aggregated for each object in turn. These are obtained by averaging for each object the numbers obtained for all images that are taken for that particular item. This way, there is one outcome per object with respect to the prediction and one output from the XRF, which resembles the real practice. The test results from the unfrozen and retrained model are used for this computation. As it can be observed, the outputs are enhanced for all the used metrics as compared to those on the images. This is due to the fact that there may be images of the same object that expose a certain degree of noise during the acquisition, or corrosion is more present.

Figure 4 indicates the manner in which the model made the prediction for each test item in part versus the XRF approximation. The objects are ordered by XRF estimation. It can be observed that the errors are very small, and the only object where the XRF value is outside of the shadow is number 6. The model performs similarly well for the objects that contain high or low percentages of Fe.

4.2. Experiment 2: Approximation of the concentration of Cu

The same data set, cross-validation splits, architecture and parameter tuning are kept for this regression instance.

4.2.1. Results and visualization

Figure 5 illustrates the wide spread of the concentration of Cu among the training/validation and test samples. Here, the distribution is shown over three intervals, as Cu is part of an alloy composition.

Table 3 shows the average results over the 5 runs for validation and test, recording the same metrics discussed previously.

Figure 6 shows in the first plot the XRF approximation against the DL prediction for the Cu samples, computed in mean over the 5 runs of the model on the test set. The second plot shows the comparison with respect to objects this time, when the value predicted for an item is the average over the predictions for its microscopic images.

4.2.2. Discussion

Figure 5 outlines the distribution of the Cu concentration over the examples in the data set. The values of Cu have a more balanced spread, which, as stated before, is normal since the metal is part of alloys in the bronze artefacts. However, this wide span makes the regression task more difficult.

The regression results for Cu are presented in Table 3. For this problem instance, the difference between the performance on validation versus test is large. The results are in general clearly better for Cu with respect to all

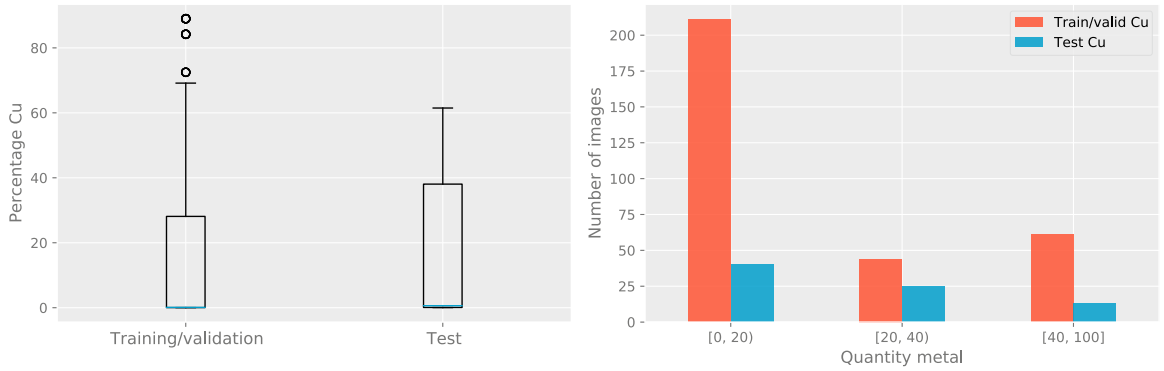

Fig. 5. The left graphic shows box plots with the percentage of Cu for the samples in the training and validation sets, as used in the 5-fold cross-validation, and in the test set. The right plot shows the number of images that contain Cu in concentrations arranged in three intervals.

Table 3. Mean validation/test metrics obtained by the model after the first stage of training and the subsequent one on recognizing the Cu amount.

Set & variant	MSE	MAE	MedAE
Validation frozen	59.59	4.69	2.16
Validation unfrozen	56.11	4.55	2.2
Test frozen	213.98	9.4	3.87
Test unfrozen & retrained	187.77	8.91	4.09
Test mean over pics per object	154.67	8.54	4.57

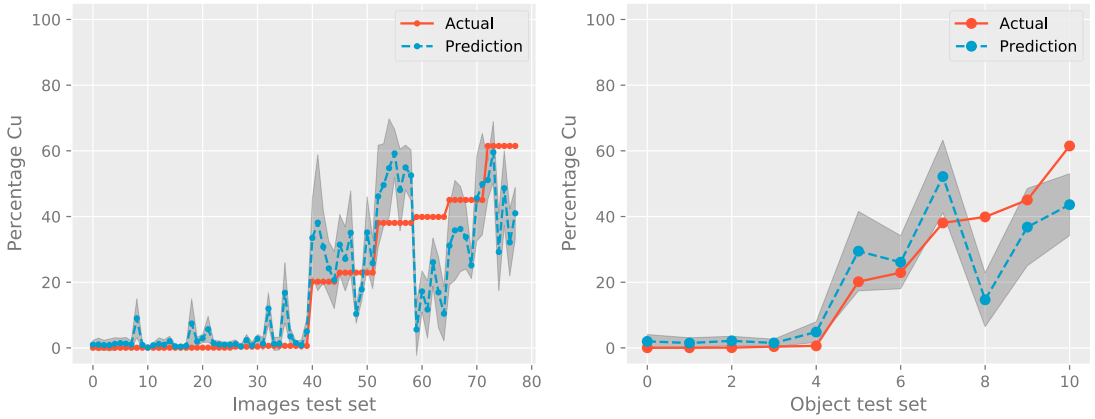

Fig. 6. XRF values for Cu (red), and average predictions (in blue) with standard deviation as a shadow for the test images (left) and objects (right).

metrics, when the second training is made by unfreezing all the layers. The MedAE value is a lot better than the MAE, because of the presence of outliers, due again to the difference in acquired images. Similarly to the case of Fe objects, the last row of the table depicting objects shows more accurate results than those of the images.

The explanation for the large difference in quality of the results for Fe and Cu lies in the manner in which the quantity of the two metals appears in the samples from the data set. While for Fe the quantity is either above 60% or under 20%, for Cu the quantity is spread over all intervals. In these circumstances, the task becomes thus less complex for Fe, while in the case of Cu, there is need for a larger exposure to bronze objects for the DL model to tune its prediction. Still, with the limited amount of data available for the moment and given that this represents the first

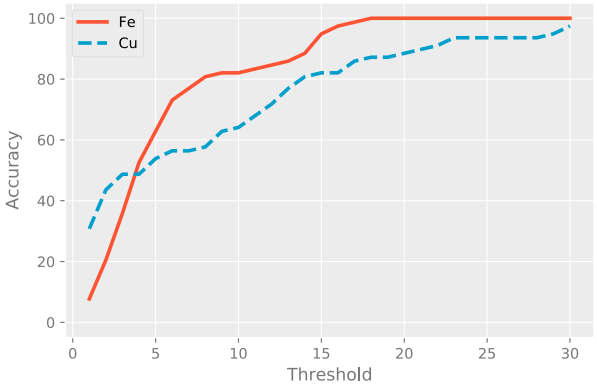

Fig. 7. The accuracy of prediction on the test set for different windows of allowed threshold.

attempt of estimating elemental composition only from microscopic images, the MAE result of about 9, as well as the MedAE that is close to 4, represent acceptable error rates for the expert in archaeology, as compared to the XRF approximation of the chemical concentrations. A deeper discussion regarding this aspect is given in subsection 4.3.

Figure 6 shows the model predictions versus the XRF output for every test image and artefact in increasing order of XRF value. Here, the errors become more prevalent with the increase in the Cu quantity, since there are less samples with such values in the data set. There are two cases for which the model heavily underestimated the amount of Cu contained in the objects, falling outside of the shadow, i.e. 8 and 10. The learning behaviour of the model is explained by the fact that the number of samples is not sufficient for the large span of the Cu concentrations in the alloy items.

4.3. Experiment 3: Prediction with an error window

The paper [2] states the limitations of the XRF technology in estimating the composition at the surface of copper alloy artefacts, due to the presence of corrosion, inherent from time and interaction with the soil. In this light, since even the result of the XRF is only approximate, the prediction of the DL model learning from that as ground truth should also be treated within a given window of acceptability.

The study [14] dealing with age prediction introduces a new metric called accuracy@t, as the percentage accuracy of the prediction given an error window of $\pm t$. We also adopt this approach, deeming it suitable for the current real-world task. Since the value of the XRF itself is primarily treated as an approximation in restoration, then implicitly the prediction of the DL learner, which is moreover based on the XRF value as a ground truth, should be also expected within a given range rather than as a singular absolute value. Keeping the notations in Table 1, this implies that, given a threshold t, if $|\hat{y}_i - y_i| < t$ then the prediction $\hat{y}_i$ for sample i is considered similar to the XRF approximation y_i.

Figure 7 shows the change in the value of the test accuracy when the threshold t is varied from a minimum of 1 to a maximum of 30. It can be seen that at $t = 7$, the accuracy for the detection of Fe already goes beyond 80%, while for Cu it needs a slack window of a double value, i.e. $t = 14$. Translated into an example, this means that if the prognosis of the DL for Cu in an image is 24, then the slack output given to the restorer is a concentration estimated in the interval [10, 38]. The accuracy for Fe already reaches 100% at $t = 17$.

5. Conclusions

The corroboration of XRF data with stereo microscopic digital images was used to train a primary deep learning regression model that would further assess the percentage of a given element in heritage objects in a completely independent way. Two problem instances were considered, one for determining the concentration of a predominant metal (Fe) in an object and another, harder instance, for assessing the quantity of an element from an alloy composition (Cu). The Fe test results are similar to those of the XRF scanner and the Cu prediction comes at an approximation

deemed as acceptable by the restorers. The processing is done in a time-efficient manner, at no additional cost apart from the stereo microscopy scan and allowing a human-free operation.

The subsequent step is to determine simultaneously all the main chemical elements on the surface of the object and thus reach a framework that would entirely mimic the XRF technology. Together with a computational estimation of corrosion [15, 16], it would then effectively help restorers towards a rapid, easy, cheap and non-invasive identification and understanding of the state of conservation of the artefacts and undertake the best subsequent procedures.

With the gathering of more objects by the restorers of the Oltenia Museum, an increased number of images will be reached and the results for Cu are expected to improve, due to the fact that a higher number of samples with different Cu concentrations in the alloy can be learnt, as this was shown to be a limitation of the current data set.

Acknowledgement

This work was supported by a grant of the Romanian Ministry of Research and Innovation, CCCDI – UEFISCDI, project number 178PCE/2021, PN-III-P4-ID-PCE-2020-0788, *Object PErception and Reconstruction with deep neural Architectures* (*OPERA*), within PNCDI III, by the Spanish Ministry of Science and Innovation, through the Plan Estatal de Investigación Científica y Técnica y de Innovación, Project TIN2017-88728-C2-1-R, as well as the *Plan Propio de Investigación, Transferencia y Divulgación Científica* of the Universidad de Málaga.

References

[1] S. A. Lins, E. Di Francia, S. Grassini, G. Gigante, S. Ridolfi, Ma-xrf measurement for corrosion assessment on bronze artefacts, in: IMEKO TC-4 International Conference on Metrology for Archaeology and Cultural Heritage, 2019, pp. 538–542.

[2] H. W. Nørgaard, Portable xrf on prehistoric bronze artefacts: Limitations and use for the detection of bronze age metal workshops, Open Archaeology 3 (1) (2017) 101–122.

[3] A. Brunetti, J. Fabian, C. W. L. Torre, N. Schiavon, A combined xrf/monte carlo simulation study of multilayered peruvian metal artifacts from the tomb of the priestess of chornancap, Applied Physics A 122.

[4] A. Rakotondrajoa, M. Radtke, Machine learning based quantification of synchrotron radiation-induced x-ray fluorescence measurements - a case study, Machine Learning: Science and Technology 2.

[5] E. Charalambous, M. Dikomitou-Eliadou, G. M. Milis, G. Mitsis, D. G. Eliades, An experimental design for the classification of archaeological ceramic data from cyprus, and the tracing of inter-class relationships, Journal of Archaeological Science: Reports 7 (2016) 465–471.

[6] E. Pouyet, J. Delaney, K. Brummel, S. Webster-Cook, C. Dejoie, G. Pastorelli, M. Walton, New insights into pablo picasso's la miséreuse accroupie (barcelona, 1902) using x-ray fluorescence imaging and reflectance spectroscopies combined with micro-analyses of samples, SN Applied Sciences 2 (2020) 1408.

[7] O. Cameron, Using deep learning to predict steering angles (2016).
URL https://medium.com/udacity/challenge-2-using-deep-learning-to-predictsteering-angles-f42004a36ff3

[8] N. Ilyas, A. Shahzad, K. Kim, Convolutional-neural network-based image crowd counting: Review, categorization, analysis, and performance evaluation, Sensors 20 (1).

[9] J.-H. Lee, S. Sull, Regression tree cnn for estimation of ground sampling distance based on floating-point representation, Remote Sensing 11 (19).

[10] A. Dobrescu, M. V. Giuffrida, S. A. Tsaftaris, Understanding deep neural networks for regression in leaf counting, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 2600–2608.

[11] B. Abirami, T. Subashini, V. Mahavaishnavi, Gender and age prediction from real time facial images using cnn, Materials Today: Proceedings 33 (2020) 4708–4712, international Conference on Nanotechnology: Ideas, Innovation and Industries.

[12] G. Levi, T. Hassncer, Age and gender classification using convolutional neural networks, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015, pp. 34–42.

[13] L. N. Smith, A disciplined approach to neural network hyper-parameters: Part 1 – learning rate, batch size, momentum, and weight decay (2018). arXiv:1803.09820.

[14] A. Smith, M. Gaur, What's my age?: Predicting twitter user's age using influential friend network and dbpedia (2018). arXiv:1804.03362.

[15] A. Samide, C. Stoean, R. Stoean, Surface study of inhibitor films formed by polyvinyl alcohol and silver nanoparticles on stainless steel in hydrochloric acid solution using convolutional neural networks, Applied Surface Science 475 (2019) 1 – 5.

[16] A. Samide, R. Stoean, C. Stoean, B. Tutunaru, R. Grecu, Investigation of polymer coatings formed by polyvinyl alcohol and silver nanoparticles on copper surface in acid medium by means of deep convolutional neural networks, Coatings 9 (2019) 105.