

Article

GRaVN: A Convolutional Network Approach to Generalised Characterisation of Raman Spectra for Space Exploration

Jon Kissi ^{1,3,†}, Tianqi Xie ^{2,3,‡}, Ken McIsaac ^{1,3,†}, Gordon R. Osinski ^{2,3,‡}, Sean Shieh ^{2,3,‡}

¹ Department of Electrical and Computer Engineering, Thompson Engineering Building, Western University, London, Ontario, Canada, N6A 5B9

² Department of Earth Sciences, Western University, 1151 Richmond Street N., London, Ontario, Canada, N6A 5B7

³ Institute for Earth and Space Exploration, Western University, 1151 Richmond Street, London, Ontario, Canada, N6A 3K7

* Correspondence: jkissiam@uwo.ca

† Current address: Affiliation 1, 3

‡ Current address: Affiliation 2, 3

Abstract: During planetary exploration mission operations, one of the key responsibilities of the instrument teams to determine data viability for subsequent analysis. During the 2019 CanMoon Lunar Sample Return Analogue Mission, the Lead Raman Specialist manually examined each spectra to provide quality assurance/validation. This non-trivial process requires years of experience to complete accurately. With the proven efficacy of Convolutional Neural Networks (CNNs) in classification tasks, and the increased use of automation and control loops on planetary space platforms for navigation and science targeting, there is an opportunity to approach this validation problem utilising CNNs. We present the Generalised Raman Validation Network (GRaVN), a neural network focussed specifically on extracting the generalised structure of Raman spectra for quality assurance/validation. This work demonstrates the viability of utilising a CNN network in validation activities for Raman spectroscopy. Utilising only two hidden layers, a configuration was developed that provided good levels of accuracy on a manually curated dataset. This indicates that such a system could be useful as part of an autonomous control loop during planetary exploration activities.

Keywords: GRaVN; machine learning; convolutional neural networks; CNN; raman spectroscopy; analogue missions; planetary science; random undersampling; random oversampling; CanMoon



1. Introduction

Citation: Kissi, J.; Xie, T.; McIsaac, K.; Osinski, G. Title. *Preprints* **2021**, *1*, 0. <https://doi.org/>

Received:

Accepted:

Published:

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

One of the key tasks for science teams on planetary surface missions is data acquisition and the activities associated with the processing of that data. In general, the initial focus is on viability, checking that the data is useful, before examining it for science. Raman spectroscopy in particular, requires a significant amount of experience and training to interpret returned data.

The focus of this work is to use machine learning to reliably automate this qualitative step, so that the remote platform (robotic rover/lander etc) itself can make decisions about data quality. When the platform takes a reading using its on-board Raman instrument, if it can detect whether it has an over-saturation (fluorescence) spectrum, for example, then it can immediately retake the reading until it obtains a qualitatively acceptable measurement.

This simple feedback loop can then be incorporated into existing sequence automation routines, already tested as a progression from manual sequencing in previous analogue missions [1–7]. The challenge with Raman spectra is that an (un)usable spectra can have a very large variety of forms. This is why we have elected to focus on extremely generalised categorisation, that the trained human conducts, rather than attempting to train on individual effects visible in unusable spectra, such as cosmic rays or fluorescence spectra, or, identify specific spectra of materials as is done for classification tasks.

1.1. Background: Raman Spectroscopy

Raman spectroscopy is a non-destructive spectra-based sensing technique, with applications in material and organic science [8,9]. In recent years, instruments specialising in the technique have seen consistent use in geological analysis terrestrially, as well as, on ground-based space platforms [10,11]. Increased use in a variety of other areas such as medical diagnosis and analysis, and, crime scene investigation [12–16], have demonstrated viability as a reliable sensing paradigm. Most recently, the Raman technique is represented on the Mars 2020 rover, Perseverance, through the SuperCam SHERLOC instrument [10]. This instrument utilises 523nm wavelength to take readings and differs from other instruments by reliably being able to take readings from further away (anything up to 7m).

Raman spectroscopy exploits reactive shifts in the frequency of monochromatic light upon interaction with [molecular species in] target materials. *Inelastic* scattering occurs when monochromatic laser light of wavelength λ_0 and energy $E_0 = hc/\lambda_0$ is incident on a surface, where h is Planck's constant and c is the velocity of light [17]. In this process, the light that interacts with the matter will either be scattered or absorbed. Of the scattered light, the majority will scatter elastically, in a process known as *Rayleigh Scattering*. However, a proportionally small amount will scatter *inelastically*, producing an emission "Raman spectrum" of light, 6 - 8 orders of magnitude weaker than the source [17,18].

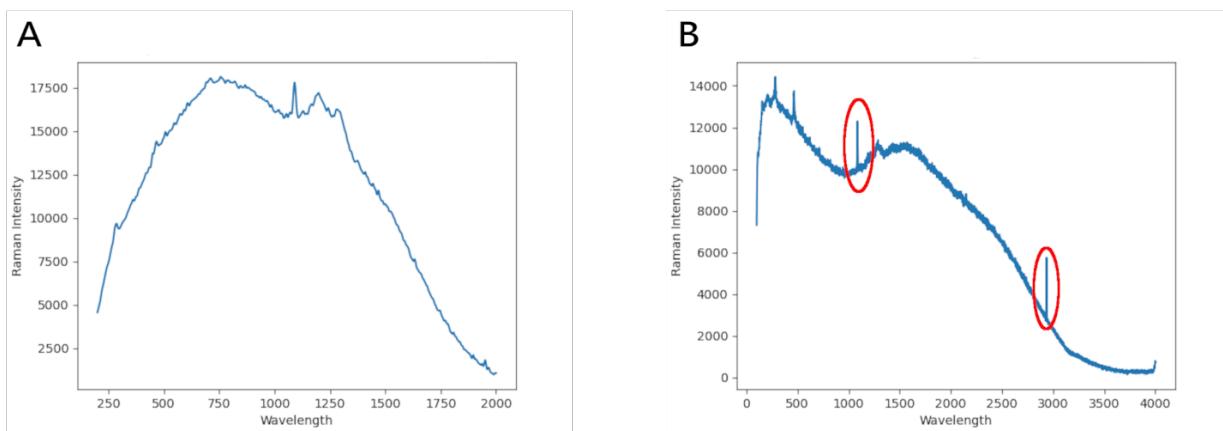


Figure 1. Raman Spectrographic Artifacts. Raman spectroscopy is sensitive to a number of factors that can prevent the acquisition of scientifically viable readings. Two of the conditions that can require retaking of the sample or post-processing are Fluorescence (A) and Cosmic Ray Spikes (B) [19–21]

1.1.1. Fluorescence Background and other Impacts to Data Quality

One of the major obstacles to obtaining well-defined spectra is a condition known as *fluorescence background* [12,18,20–26]. This has the effect of "over-saturating" the spectra, such that the identifying molecular peaks are no longer visible. Many molecules that respond to Raman excitation can also exhibit fluorescence behaviour upon excitation, so this fluorescence state is frequently encountered in complex samples [22], which is unfortunate, because *apparent* complexity, particularly in a planetary exploration situation, is often the reason that a target is selected to start with. Fluorescence can also have

the knock-on effect of increasing *shot noise*, a saturation effect associated with Rayleigh scattering [8,9,27], further adding to the difficulty of validation.

A variety of hardware and software-based methods have been explored for quenching fluorescence background and enhancing weak Raman signals. These range from modifying the aperture time of the Raman instrumentation, so that there isn't enough time for the fluorescence to develop in the signal [12,18,20–26], to applying baseline detection algorithm to estimate a clean signal [28]. All these techniques interfere with the data itself by either modifying conditions for acquisition or altering data directly, which impacts interpretation.

Another artifact that can appear in spectra results are Cosmic Ray spikes. These manifest as sharp spikes that often extend vertically by a disproportionate amount. These spikes are representative of low level radiation events and appear randomly across different areas of the spectrum [9]. Generally the most used technique for removal of these, is to compare spectra of the same target and remove spikes that are not contained in both. [8,9]

The third major consideration when evaluating Raman return signals is the signal-to-noise (SNR) ratio present in the reading. In general the lower signal-to-noise, the better quality the reading. Various mathematical and algorithmic methods have been utilised to improve SNR including derivative and wavelet methods [20], Fourier filtering and least squares analysis [21]. In this regard some of the best results came from the use of Prime Component Analysis, where the authors demonstrate that this technique reduced acquisition times, allowing for it to be used in real-time applications [21].

1.2. Raman Technique Compatibility with Automation and Machine Learning

For commonly used spectra-based techniques such as Laser Induced Breakdown Spectroscopy (LIBS) [29], Alpha Particle X-Ray Spectrometer (APXS) and Mass Spectroscopy for example, a peak at a specific wavelength will usually indicate a specific material substance and the result can be reproduced reliably, regardless of mitigating factors [30,31]. For Raman, that is not necessarily the case. Here, the same species or molecule can have multiple signatures at different positions. This aspect is what adds an extra challenge to the automated classification of Raman data.

This, plus the results of various conditions under which that sample has been captured, lead to an intrinsic multi-variate uncertainty that suggests a complex problem space. The increasing application of machine learning in planetary science and space exploration for problems residing in complex problem spaces [4,7,16,28,29,32–41], suggests an opportunity to further apply these techniques.

1.3. Raman Analysis for Analogue and Live Planetary Rover Missions

Terrestrial analogues are locations on Earth that approximate conditions on other planetary bodies. These locations allow for comparative planetology studies, technology development/demonstration and the training of personnel.

Analogue missions are integrated field campaigns, conducted in terrestrial analogue environments. The 2019 CanMoon [1,3,6,42–44] formed part of the Canadian Space Agency's Lunar Exploration Analogue Deployment program, with requirements to:

1. Compare the accuracy of selecting lunar samples remotely from mission control versus a traditional human field party
2. Test the efficiency of remote science operations including the use of pre-planned strategic traverses
3. Evaluate the utility of real-time automated data analysis approaches for lunar missions
4. Explore the mission control operations structure for 24/7 lunar science operations

5. Test how Virtual Reality (VR) technology can be used to help with enhancing the situational awareness in mission control

In order to achieve this, a full complement of instrument types was deployed in the field, with science and navigation targeting controlled from Mission Control at Western University, over the course of several days [1,6,42]. During the mission, one of the key responsibilities of the Raman specialist was to determine data viability for analysis [7]. The lead Raman specialist was a geoscience PhD student who had been training to recognise Raman data from a variety of instruments and conditions. Every spectra returned from targeting activities was manually evaluated in order to provide a "Good", "Bad" or "Potential" categorisation. This indicated whether the data met the criteria to attempt a detailed scientific classification. This task was non-trivial, requiring significant time on the downlink end of operations to complete. It could not be completed automatically, or via a well-defined pre-mission training process i.e. during operations time because of the complexity of analysis. Amongst the Raman instrument team, there was often vigorous debate as to whether the data received contained a viable spectra, indicating the challenge in interpreting results.

Table 1: Breakdown of samples per category by origin

Source	Good	Potential	Bad
RRUFF	2597	730	198
Lab	42	56	28
Handheld	3	11	5
Other	13	0	8

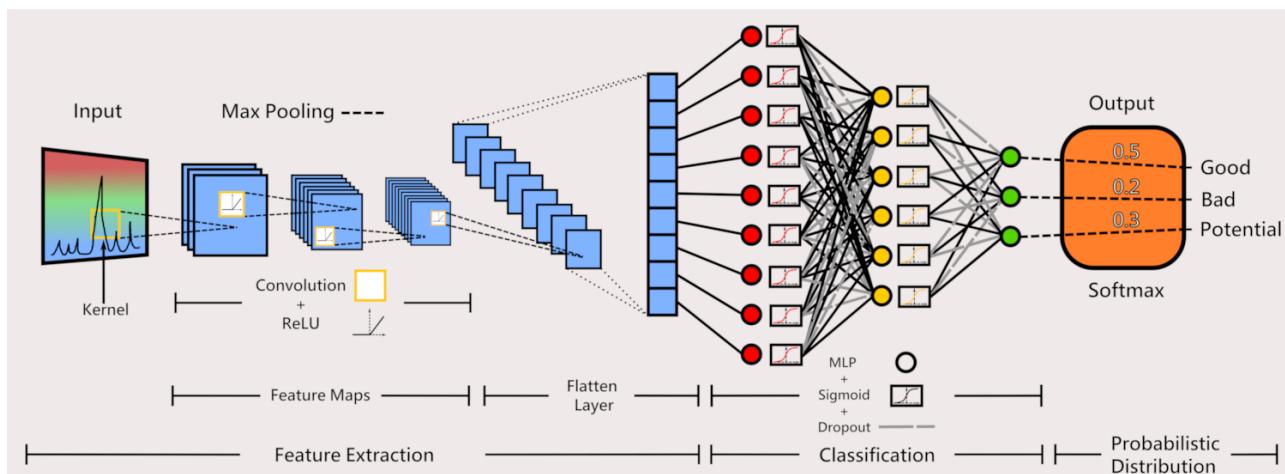


Figure 2. The GRaVN Convolutional Network. The network consists of 3 2D Convolutional layers 16, 32, 64 neurons deep respectively, with ReLU activation, with 2D max pooling. Following are 2 fully connected layers of neurons with Sigmoid activation, dropout of 0.5 and a final classification layer with Softmax. Sparse Categorical Cross Entropy (SCCE) was used as the model loss function, with the ADAM algorithm used as the optimiser, initialised with a learning rate of $10e^{-4}$.

In the case of Raman data validation, to be competent takes skill, patience and experience. Also with the requirement of analogue missions to maintain fidelity to real planetary surface operations, data can only be reviewed intermittently (during communication cycles), and bandwidth is limited [2,5,42,45]. So, in this case, the "best" option is to send *all* collected data back for further processing and have a trained technician perform data validation before then continuing on to extract science. This is not optimal, and makes it likely that at least some resources will be wasted during the communication cycle.

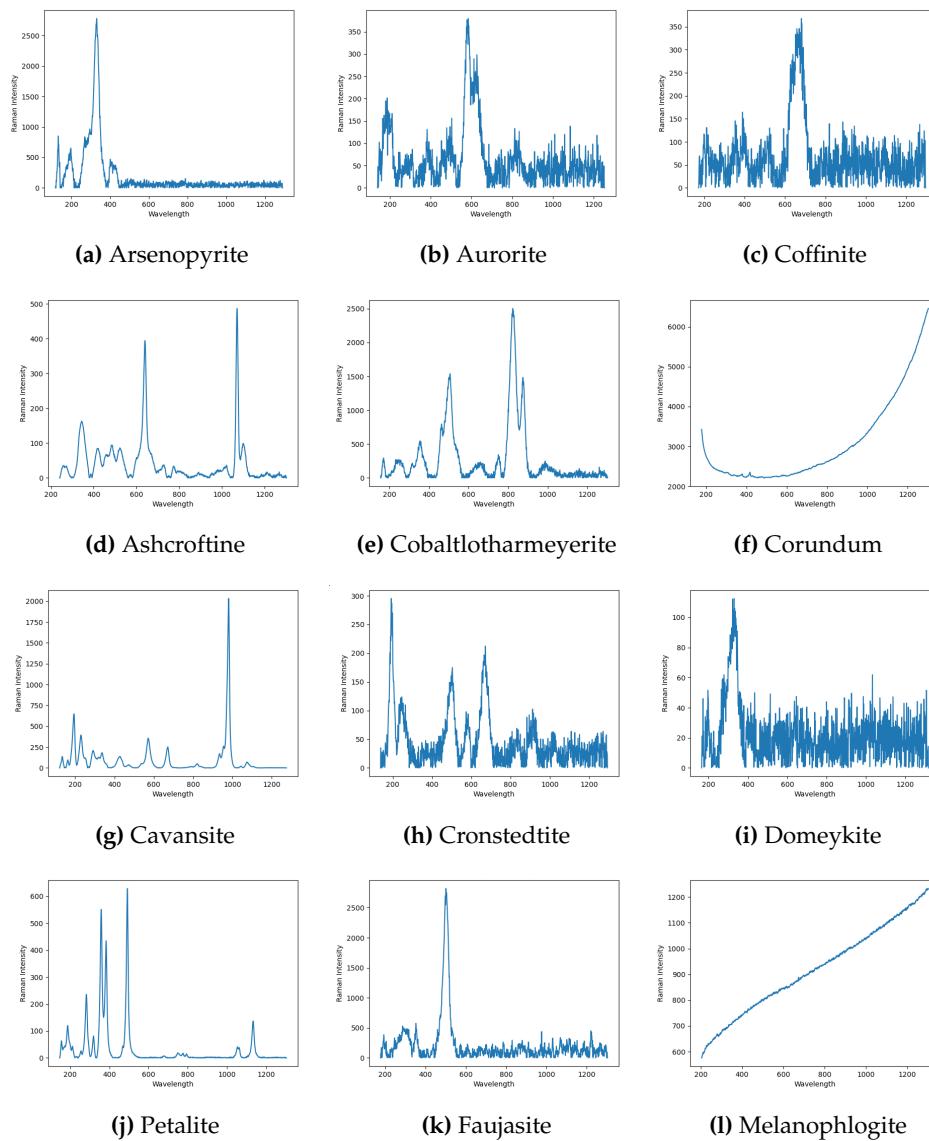
1.4. Algorithmic Processing for Raman Data

Work in this area revolves around using numerical and computational techniques to treat data, so that all the data returned can be considered viable. Here, a variety of work has been done in the area of fluorescence suppression [21,22,25,26], spectra identification using standard algorithmic processing [15,24,28,46] and more advanced techniques, such as machine learning and/or CNNs [33,34,46–49], to directly identify the molecular signatures.

However, there are limitations to this. Some spectra cannot be improved by these techniques, and there is the technical point that these techniques *alter* the original data, and we believe there is significant merit in preserving the original form of the data, before it is evaluated by a trained technician.

Taking into account these points, plus the sensitivity of the readings to additional conditions not related to the physical construction of the instrument, an argument can be made that the option that is likely to result in the best data analysis, is to *start* with the best spectra possible.

Thus, a reliable machine learning model, that can perform initial data validation in the same way as an experienced technician, could prove invaluable in saving resources. By providing in-situ analysis of quality as data is acquired, an automated command cycle can be implemented, with the aim that *only* valid data is returned for evaluation.



1.5. Neural Network Approaches to Raman Analysis

There are many works in the literature focussing on identification of specific spectral signatures, in a variety of different fields [17,19,27,32,46,49]. The focus of this work is not to propose another solution to that problem but instead, to focus on **initial quality assurance/validation**, based on *generalised criterion*.

While the criterion is more general, the act of categorisation is still a non-trivial operation, particularly because the data conditions, such as the wide variance of SNR in samples, make it difficult to simply cross-check with a database, for example.

As mentioned previously, this can only be done efficiently once the **human technician** has gone through the process many times. In order for the technician to become competent, the process must be completed using multiple different instruments, instrument types (field, laboratory, handheld, etc), and different conditions under which readings are taken. All these variables can affect the resulting spectral signature of the data products produced, even if the target is the same [9].

Recognition of all these elements feeds into the expert technicians' decision as to whether a data product can be considered contain data of high quality (*Good*), unusable data (*Bad*) or data that can be further treated to produce usable data (*Potential*).

This kind of multi-variate problem, where a "correct" answer is not easily defined, is generally considered a space to which neural networks are well suited. This work will examine the viability of a simple, deep 2D Convolutional Neural Network in performing this *generalised categorisation* on Raman Spectra.

2. Materials and Methods

2.1. Software Pipeline

The GRaVN pipeline reads each comma separated datafile and generates a 300 x 300 pixel spectra for consumption by the GRaVN network. The larger than usual image size is to allow enough base resolution to reliably pick up detail at small and large scales. Once this operation is complete, each spectra is converted to grey scale and matched with the appropriate label before being preprocessed, read to pass into the CNN. More information in the preprocess step is given in section 2.4.1 and information on the neural network itself is given in section 2.2.

2.2. Network Architecture and configuration

The Convolutional network consists of a layer configuration of three 2D Convolutional layers 16, 32, 64 neurons deep respectively, each with a window size of 3x3 and ReLu activation. This is then followed by 2D max pooling layer with a pool size of 2x2. The feature maps are flattened and then passed to two fully connected layers. The first consists of 1024 neurons and an l2 regulariser. The activation is a sigmoid function and dropout of 0.5 is applied to the layer. The following layer utilises the same configuration but with 256 neurons. The final layer is a fully connected output layer consisting of a number of neurons equalling the total number of classes, with a standard softmax activation. Sparse Categorical Cross Entropy was used as the model loss function, with the ADAM algorithm used as the optimiser. See Fig. 3

2.3. The Dataset

The dataset is manually curated from multiple sources by a trained PhD-level Raman specialist. It consists of a total of 3691 spectra, labelled as either Good, Bad, Potential. The breakdown between the categories is given in Table 1. The dataset consists of approximately

3500 samples taken from the 532 nm catalog in the RRUFF database, with the remainder being manually curated from other sources, such as lab-based and handheld spectrometers such as the DeltaNu RockHound.

2.4. Data Preprocessing

As poor quality readings aren't generally considered useful, it is far more common to find examples of good spectra than any other. For the purpose of training a CNN-derived model, designed to sort the spectra into different classifications, this leads to what is referred to as an *imbalanced dataset*. The problem of imbalanced datasets is a common

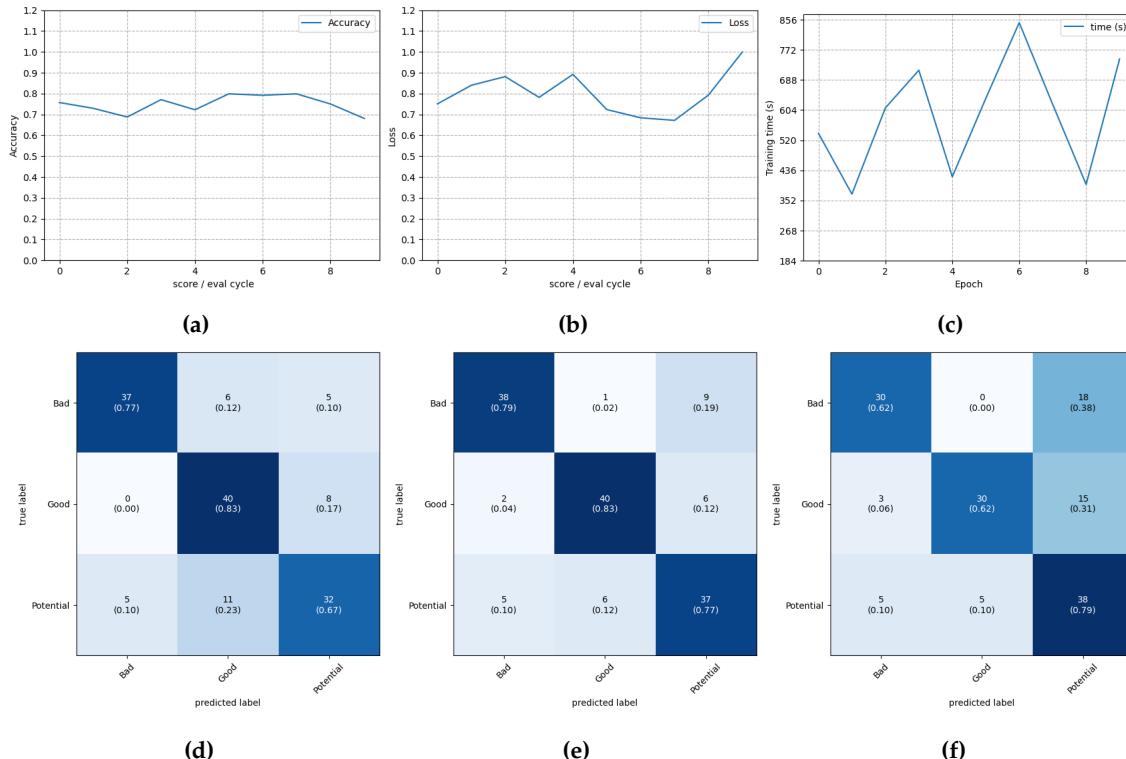
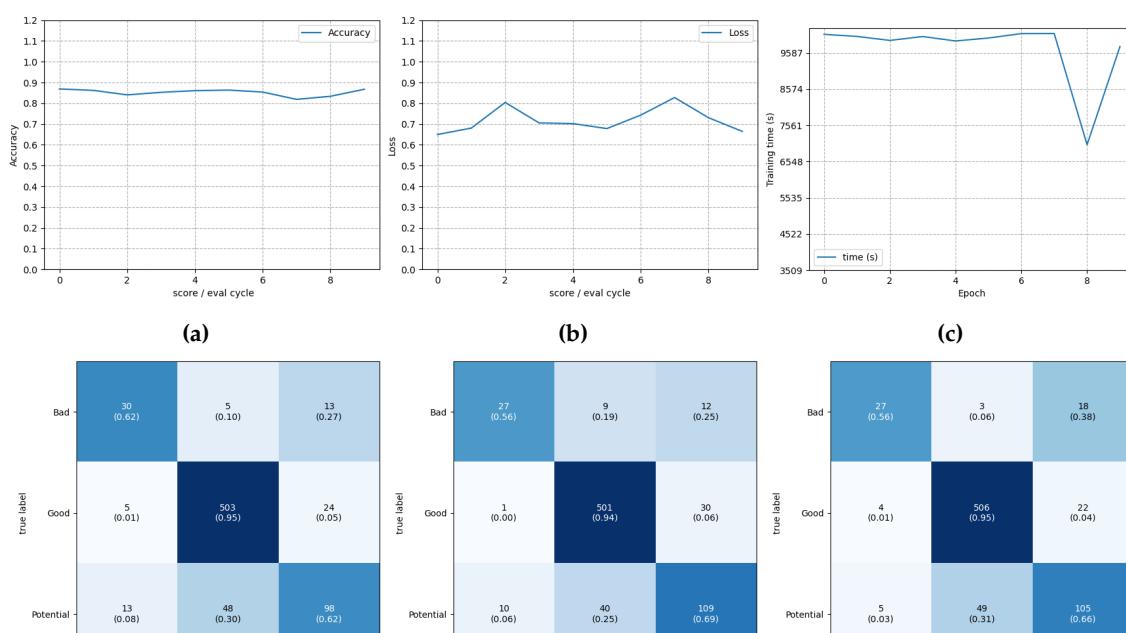


Figure 4. Results of the GRaVN network with a RU strategy. Top left to top right: (a) Accuracy, (b) Loss and (c) training/validation time to completion for each epoch. Bottom left to bottom right: Confusion Matrices for evaluation cycles 1 (e), 5 (f), 10 (g), respectively.



one in machine learning and there are several strategies to deal with this. Mitigation techniques are discussed in section 2.4.1.

2.4.1. Sampling Methods

We chose to employ random sampling to smooth network learning between categories. In this case, we chose two techniques: Random Undersampling (RU) and Random Over-sampling (RO). These strategies are commonly accepted substitutes for not having a large, balanced dataset [50–52] but both have trade-offs in implementation: RU limiting variation in training, the larger the class is, and, RO potentially oversampling from smaller classes. Results from both are examined to determine which, if any, work better for our case.

2.5. Experiments and Results Validation

We tested the network using a variation on the 10-Fold Cross Validation methodology. In our case, each “fold” is an experimental run consisting of a training and validation cycle, followed by a testing (evaluation) cycle on unlabelled data. Because of the limitations of the size of our dataset, each run uses the *entire* dataset. The dataset is split 80%/20% train/test, with the same split on the training set for train/validation.

For each RO run the train test split occurs immediately followed by the sampling operation. This ensures that the train set contains no spectra that are in the test set.

For RU, the sampling operation occurs first, followed by the train-test split. This, again, ensures that no evaluation spectra are contained in the training set. Numbers of samples for each category are contained in Table 1.

Experiments consisted of standard evaluation of accuracy, loss and time to completion metrics for training/validation, and, accuracy and loss for evaluation. Training cycles were set for a maximum of 30 epochs, with early stopping activated after no accuracy score improvement for 5 epochs.

Table 2: Statistical table showing results from Confusion matrices for each class.

Sampling Type	Min	Max	Median	Mean	σ (std)	σ^2 (variance)
Undersampling						
Good	0.620	0.960	0.830	0.828	0.0991	0.0098
Bad	0.620	0.810	0.730	0.720	0.074	0.006
Potential	0.560	0.850	0.700	0.695	0.098	0.009694
Oversampling						
Good	0.920	0.950	0.940	0.938	0.010	0.0001
Bad	0.460	0.710	0.580	0.593	0.080	0.006
Potential	0.500	0.690	0.675	0.646	0.063	0.004

3. Results

For RU, figure 5 shows accuracy at evaluation time consistently remains between 70% and 80% approx. for each run. Loss tends to fluctuate more, between approximately 0.68 and 1.0. Training time varied between approximately 350 and 850 seconds for each run.

Conversely the RO results fluctuate far less. Accuracy remains in a tighter value band between 80% and 90%, while loss remains between approximately 0.65 and 0.8. Time to

completion for each run, was much longer, however ranging from well over 9500 to just over 6500 seconds, which of course makes sense, considering there is a larger overall size of the training set.

When comparing Confusion Matrices overall (Table 2), RO and RU have almost the exact same std and variance for Good and Bad classes. Oversampling is better for the Potential group. The average correct classification score for Good is higher for oversampling while for Bad and Potential classes undersampling averages better.

Table 3 shows that for accuracy and loss, the RO average score is higher and lower for accuracy and loss, respectively, suggesting better performance than with a RU treated dataset. In this case, the lower std and variance for RU confirm the lower rate of variation for RO shown in Figure 6, as compared to RU in Figure 5.

During evaluation, the GRaVN model classifies with accuracy scores consistently above 70% for both sampling techniques. That is, post training cycle, on unseen data the CNN is guaranteed to be correct 70% of the time. We contend that this demonstrates the robustness of the GRaVN model for this task. RO showed the best results for evaluation accuracy, with scores consistently between 80%-90% as opposed to 70%-80%.

4. Discussion

In the case of RO, the test set is extracted before the sampling operation, then the all elements of the smaller classes are randomly duplicated until all classes have the same number as the largest class. What these results indicate is that for RO the network learns the Good class with a high level of accuracy, but because of the proportional difference between the class sizes the network, will over train on the smaller classes. This means that when presented with the test/evaluation set, that the network finds it difficult to accurately identify the Bad classes in particular, while showing a high level of accuracy when classifying a Good spectra.

For RU, the result is slightly different. In this case, to preserve the integrity of the test set, the sampling operation is performed first, then the 80%/20% train/test split is done. This means that training occurs on even numbers of each class but those numbers are smaller, resulting in lower accuracy scores overall, but a more even distribution of accuracy between the classes.

Table 3: Statistical table showing results from Evaluation Accuracy, Loss and time to train the model.

Sampling Type	Min	Max	Median	Mean	σ (std)	σ^2 (variance)
Undersampling						
Accuracy	0.681	0.799	0.753	0.749	0.043	0.002
Loss	0.671	0.999	0.787	0.802	0.103	0.011
Time (s)	369	847	615	590	159	25332
Oversampling						
Accuracy	0.819	0.869	0.857	0.852	0.016	0.000268
Loss	0.649	0.828	0.704	0.718	0.059	0.003459
Time (s)	7018	10131	10025	9711	952	907728

We consider this level of generalisation challenging for a network to learn, especially with the highly imbalanced nature of the dataset. This assessment also considers the nature of the data being evaluated. A 2D convolutional network examines images pixel by pixel. When examining images of spectra like this, the vast majority of the information is contained in a small area. Within this area is a very large variation because not only can

the SNR vary, but so can the geometry and amplitude because we are not limiting training to specific signatures.

These factors create a broad problem space on which to attempt to apply solutions. Therefore, we consider this result positive and good platform to build upon. We intend to continue modifying our neural network layer depth and width, to further to improve our accuracy and loss scores. Although, because of the nature of the problem, we would not necessarily predict that a simple increase in network depth and/or width (and therefore increase in the number of trainable parameters), would guarantee an improvement in network output statistics.

He et al. first recorded that simply stacking increasing numbers of convolutional blocks in network architectures limits optimisation, and leads to diminishing returns in accuracy through overfitting and other artifacts such as the degradation problem [53]. Skip connections, which allow network information to skip layers, rather than passing directly between them, were devised as a method of maintaining the trainability of models, while increasing their depth. This would be an interesting avenue to pursue in increasing network complexity, while continuing to improve accuracy.

Further, in the area of mineral classification, certain network architectures [32] have demonstrated promising results by passing different parts of the input data through the network separately, and concatenating these in into a single dense layer for classification later in the process. A network developed utilising these tools, along with a larger dataset, with higher proportion of Bad spectra, could learn to differentiate far more of the characteristics of a Raman spectra, and thus be more accurate.

Lastly, it would be interesting to compare our network to some well-known convolutional frameworks. The closest problem to our work would be handwriting recognition. Although at the pixel scale, there is more variation in this problem because of the broad range of SNR, a factor that rarely plays a role in handwriting or letter/digit recognition. A comparison between well-known networks like VGG or ResNet and our method, could be informative in a variety of different areas.

Table 4: Statistical table showing computational process time per sample at evaluation stage.

Sampling type	Num Samples	Mean Eval Time (s)	Time/Sample (s)
RU	144	1.97	0.01368
RO	737	9.71	0.01318

5. Conclusion and Future Work

We have presented a convolutional neural network solution for Raman spectrum processing, focusing on qualitative validation rather than specific classification. The network was able to accurately sort untreated spectra into categories as specified by a trained technical expert in the field.

The results indicate that networks developed to perform the more qualitative evaluative tasks are a viable option and, that further development to improve output scores is a legitimate direction to pursue.

We suggest that an algorithm such as this could be the basis of a sequence control loop. Automated space platforms could use to validate and pre-select data in-situ during planetary missions, or other autonomous robotic activities, where Raman is utilised.

Regarding the sampling techniques for training, if mission parameters would specify a two class spectra evaluation, then the RO technique would be a better choice. In this case, mission operators would only be interested in returning spectra classed as Good. If mission

parameters dictated that a more granular validation be preferable, then and RU training scheme could be implemented. Table 4 shows that network process time per sample on the order of tens of milliseconds, regardless of sampling type and indicates the time scale required for a platform to obtain a reading.

Continuing work in this area would focus on addressing the dataset imbalance issue by acquiring more examples of poorer quality spectral reading, increasing the overall size of the dataset and, perhaps even expanding the classification to different Raman sources.

We would welcome input from the Raman community in providing us with [labelled] examples of Bad or Potentially Bad/Good spectra readings, regardless of instrument or conditions. As stated above, it is easy to find examples of good readings, but this algorithm's success depends on knowing what a bad reading looks like. So please, don't discard poor data, send it to us. It would allow us to further increase robustness by including data classified from multiple experts.

Unlike with other techniques, internal instrument components, wavelengths used and sample interval can all vary between different instruments. An interesting study would be to build a network capable of equalising all these factors and provide a tool that could accurately validate regardless of any mitigating construction or environmental factors.

Table 5: Computer Specifications

Hardware	
CPU	AMD Ryzen 7 3750H w/ Radeon Vega Mobile Gfx
RAM	32GB
GPU	Nvidia GeForce GTX 1660 Ti
Software	
Lang.	Python 3.7+
NN Arch.	Tensorflow 2.0

Author Contributions: Conceptualization, Jon Kissi and Ken McIsaac; Data curation, Tianqi Xie; Formal analysis, Jon Kissi; Funding acquisition, Ken McIsaac, Gordon Osinski and Sean Shieh; Investigation, Jon Kissi; Methodology, Jon Kissi and Ken McIsaac; Software, Jon Kissi; Supervision, Ken McIsaac, Gordon Osinski and Sean Shieh; Validation, Tianqi Xie; Visualization, Jon Kissi; Writing – original draft, Jon Kissi; Writing – review & editing, Jon Kissi, Tianqi Xie, Ken McIsaac, Gordon Osinski and Sean Shieh. All authors have read and agreed to the published version of the manuscript.

Funding: Funding gratefully acknowledged from the following bodies: 2019 CanMoon Analogue Mission: Canadian Space Agency; All other funding: Natural Sciences and Engineering Research Council (NSERC).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. Code developed in this study is available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Andres, C.; Ratcliffe, K.; Newman, J.; Osinski, G.R.; Cloutis, E.; Marion, C.L.; Pilles, E.A. ROVER COMMAND SEQUENCING AND INSTRUMENT UTILIZATION DURING THE CANMOON ANALOGUE MISSION. *51st Lunar and Planetary Science Conference 2020*, pp. 51–52.

2. Osinski, G.R.; Battler, M.; Caudill, C.M.; Francis, R.; Haltigin, T.; Hipkin, V.J.; Kerrigan, M.; Pilles, E.A.; Pontefract, A.; Tornabene, L.L.; Allard, P.; Bakambu, J.N.; Balachandran, K.; Beaty, D.W.; Bednar, D.; Bina, A.; Bourassa, M.; Cao, F.; Christoffersen, P.; Choe, B.H.; Cloutis, E.; Cote, K.; Cross, M.; D'Aoust, B.; Draz, O.; Dudley, B.; Duff, S.; Dzamba, T.; Fulford, P.; Godin, E.; Goordial, J.; Grau Galofre, A.; Haid, T.; Harrington, E.; Harrison, T.; Hawkswell, J.; Hickson, D.; Hill, P.; Innis, L.; King, D.; Kissi, J.; Laughton, J.; Li, Y.; Lymer, E.; Maggiori, C.; Maloney, M.; Marion, C.L.; Maris, J.; Mcfadden, S.; McLennan, S.M.; Mittelholz, A.; Morse, Z.; Newman, J.; O'Callaghan, J.; Pascual, A.; Patel, P.; Picard, M.; Pritchard, I.; Poitras, J.T.; Ryan, C.; Sapers, H.; Silber, E.A.; Simpson, S.; Sopoco, R.; Svensson, M.; Tolometti, G.; Uribe, D.; Wilks, R.; Williford, K.H.; Xie, T.; Zylberman, W. The CanMars Mars Sample Return analogue mission. *Planetary and Space Science* **2019**, *166*, 110–130. doi:10.1016/j.pss.2018.07.011.
3. Newman, J.; Pilles, E.A.; Morse, Z.R.; Marion, C.L.; Christoffersen, P.; Hill, P.; Osinski, G.R.; Cloutis, E.; Andres, C.; Bourassa, M.; Caudill, C.M.; Cross, M.; Dicecca, A.; Doerksen, K.; Hibbard, S.; Hopkins, R.; Kollewyn, J.; Pascual, A.; Ratcliffe, K.; Roberts, A.; Rodriguez Sanchez-Vahamonde, C.; Ryan, C.; Shah, J.; Tolometti, G.; Tornabene, L.L.; Vlachopoulos, O. PLANNING TEAM OPERATIONS FOR THE CANMOON LUNAR SAMPLE RETURN ANALOGUE MISSION. *51st Lunar and Planetary Science Conference* **2020**.
4. Francis, R.; Gaines, D.; Osinski, G.R. Advanced Rover Science Autonomy Experiments in Preparation for The Mars 2020 Mission: Results from the 2016 CanMars Analogue Mission. *Lunar and Planetary Science Conference* **2017**, *XLVIII*, 2402.
5. Caudill, C.M.; Pontefract, A.J.; Osinski, G.R.; Tornabene, L.L.; Pilles, E.A.; Battler, M.; Francis, R.; Godin, E.; Grau Galofre, A.; Haltigin, T.; Hipkin, V.J.; Mittelholz, A.; Poitras, J.; Simpson, S.L.; Svensson, M.; Xie, T.; Morse, Z.R. CanMars mission Science Team operational results: Implications for operations and the sample selection process for Mars Sample Return (MSR). *Planetary and Space Science* **2019**, *172*, 43–56. doi:10.1016/j.pss.2019.04.004.
6. Morse, Z.R.; Hill, P.; Osinski, G.R.; Cloutis, E.; Marion, C.L.; Newman, J.; Caudill, C.M.; Christoffersen, P.; Pilles, E.A.; Simpson, S.L.; Tornabene, L.L.; Xie, T. CANMOON ANALOGUE MISSION TACTICAL SCIENCE TEAM: SCIENTIFIC INSTRUMENTATION AND DECISION MAKING DURING A HIGH FIDELITY REALTIME LUNAR ANALOGUE MISSION. *51st Lunar and Planetary Science Conference* **2020**, pp. 51–52.
7. Xie, T.; Osinski, G.R. The Use of Raman Spectroscopy for the 2016 CanMars MSR Analogue Mission. *47th Lunar and Planetary Science Conference* **2016**, p. Abstract 1578.
8. Bowie, B.T.; Chase, B.D.; Griffiths, P.R. Factors Affecting the Performance of Bench – Top Raman Spectrometers . Part II : Effect of Sample. *Applied Spectroscopy* **2000**, *54*.
9. Bowie, B.T.; Chase, D.B.; Griffiths, P.R. Factors affecting the performance of bench-top Raman spectrometers. Part I: instrumental effects. *Applied Spectroscopy* **2000**, *54*. doi:10.1366/0003702001949924.
10. Maurice, S.; Wiens, R.C.; Bernardi, P.; Caïs, P.; Robinson, S.; Nelson, T.; Gasnault, O.; Reess, J.M.; Deleuze, M.; Rull, F.; Manrique, J.A.; Abbaki, S.; Anderson, R.B.; André, Y.; Angel, S.M.; Arana, G.; Battault, T.; Beck, P.; Benzerara, K.; Bernard, S.; Berthias, J.P.; Beyssac, O.; Bonafois, M.; Bousquet, B.; Boutillier, M.; Cadu, A.; Castro, K.; Chapron, F.; Chide, B.; Clark, K.; Clavé, E.; Clegg, S.; Cloutis, E.; Collin, C.; Cordoba, E.C.; Cousin, A.; Dameury, J.C.; D'Anna, W.; Daydou, Y.; Debus, A.; Deflores, L.; Dehouck, E.; Delapp, D.; De Los Santos, G.; Donny, C.; Doressoundiram, A.; Dromart, G.; Dubois, B.; Dufour, A.; Dupieux, M.; Egan, M.; Ervin, J.; Fabre, C.; Fau, A.; Fischer, W.; Forni, O.; Fouchet, T.; Frydenvang, J.; Gauffre, S.; Gauthier, M.; Gharakanian, V.; Gilard, O.; Gontijo, I.; Gonzalez, R.; Granena, D.; Grotzinger, J.; Hassen-Khodja, R.; Heim, M.; Hello, Y.; Hervet, G.; Humeau, O.; Jacob, X.; Jacquinod, S.; Johnson, J.R.; Kouach, D.; Lacombe, G.; Lanza, N.; Lapauw, L.; Laserna, J.; Lasue, J.; Le Deit, L.; Le Mouélic, S.; Le Comte, E.; Lee, Q.M.; Legett, C.; Leveille, R.; Lewin, E.; Leyrat, C.; Lopez-Reyes, G.; Lorenz, R.; Lucero, B.; Madariaga, J.M.; Madsen, S.; Madsen, M.; Mangold, N.; Manni, F.; Mariscal, J.F.; Martinez-Frias, J.; Mathieu, K.; Mathon, R.; McCabe, K.P.; McConnochie, T.; McLennan, S.M.; Mekki, J.; Melikechi, N.; Meslin, P.Y.; Micheau, Y.; Michel, Y.; Michel, J.M.; Mimoun, D.; Misra, A.; Montagnac, G.; Montaron, C.; Montmessin, F.; Moros, J.; Mousset, V.; Morizet, Y.; Murdoch, N.; Newell, R.T.; Newsom, H.; Nguyen Tuong, N.; Ollila, A.M.; Orttner, G.; Oudda, L.; Pares, L.; Parisot, J.; Parot, Y.; Pérez, R.; Pheav, D.; Picot, L.; Pilleri, P.; Pilorget, C.; Pinet, P.; Pont, G.; Poulet, F.; Quantin-Nataf, C.; Quertier, B.; Rambaud, D.; Rapin, W.; Romano, P.; Roucayrol, L.; Royer, C.; Ruellan, M.; Sandoval, B.F.; Sautter, V.; Schoppers, M.J.; Schröder, S.; Seran, H.C.; Sharma, S.K.; Sobron, P.; Sodki, M.; Sournac, A.; Sridhar, V.; Standarovsky, D.; Storms, S.; Striebig, N.; Tatat, M.; Toplis, M.; Torre-Fdez, I.; Toulemont, N.; Velasco, C.; Veneranda, M.; Venhaus, D.; Virmontois, C.; Viso, M.; Willis, P.; Wong, K.W. *The SuperCam Instrument Suite on the Mars 2020 Rover: Science Objectives and Mast-Unit Description*; Vol. 217, The Author(s), 2021. doi:10.1007/s11214-021-00807-w.
11. Rodriguez, I.; Lopez-Reyes, G.; Llanos, D.; Perez, F. Automatic Raman Spectra Processing for Exomars. *Mathematics of Planet Earth* **2014**, pp. 703–707. doi:10.1007/978-3-642-32408-6.
12. Kumar, S.; Verma, T.; Mukherjee, R.; Ariese, F. Raman and infra-red microspectroscopy: towards quantitative evaluation for clinical research by ratiometric analysis. *The Royal Society of Chemistry* **2016**, pp. 1879–1900. doi:10.1039/c5cs00540j.
13. Yan, X.; Zhang, S.; Fu, H.; Qu, H. Combining convolutional neural networks and on-line Raman spectroscopy for monitoring the Cornu Caprae Hircus hydrolysis process. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* **2020**, *226*, 117589. doi:10.1016/j.saa.2019.117589.
14. Fujihara, J.; Fujita, Y.; Yamamoto, T.; Nishimoto, N.; Kimura-Kataoka, K.; Kurata, S.; Takinami, Y.; Yasuda, T.; Takeshita, H. Blood identification and discrimination between human and nonhuman blood using portable Raman spectroscopy. *International Journal of Legal Medicine* **2017**, *131*, 319–322. doi:10.1007/s00414-016-1396-2.
15. Virkler, K.; Lednev, I.K. Blood species identification for forensic purposes using Raman spectroscopy combined with advanced statistical analysis. *Analytical Chemistry* **2009**, *81*, 7773–7777. doi:10.1021/ac901350a.

16. Tarcea, N.; Harz, M.; Rösch, P.; Frosch, T.; Schmitt, M.; Thiele, H.; Hochleitner, R.; Popp, J. UV Raman spectroscopy-A technique for biological and mineralogical in situ planetary studies. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* **2007**, *68*, 1029–1035. doi:10.1016/j.saa.2007.06.051.
17. Cochrane, C.J.; Blacksberg, J. A Fast Classification Scheme in Raman Spectroscopy for the Identification of Mineral Mixtures Using a Large Database with Correlated Predictors. *IEEE Transactions on Geoscience and Remote Sensing* **2015**, *53*, 4259–4274. doi:10.1109/TGRS.2015.2394377.
18. Sevetlidis, V.; Pavlidis, G. Effective Raman spectra identification with tree-based methods. *Journal of Cultural Heritage* **2019**, *37*, 121–128. doi:10.1016/j.culher.2018.10.016.
19. Zhang, D.; Ben-Amotz, D. Enhanced chemical classification of raman images in the presence of strong fluorescence interference. *Applied Spectroscopy* **2000**, *54*, 1379–1383. doi:10.1366/0003702001951066.
20. Zhang, Z.M.; Chen, S.; Liang, Y.Z.; Liu, Z.X.; Zhang, Q.M.; Ding, L.X.; Ye, F.; Zhou, H. An intelligent background-correction algorithm for highly fluorescent samples in Raman spectroscopy. *Journal of Raman Spectroscopy* **2010**, *41*, 659–669. doi:10.1002/jrs.2500.
21. Mazilu, M.; De Luca, A.C.; Riches, A.; Herrington, C.S.; Dholakia, K. Optimal algorithm for fluorescence suppression of modulated Raman spectroscopy. *Optics Express* **2010**, *18*, 11382. doi:10.1364/oe.18.011382.
22. Wei, D.; Chen, S.; Liu, Q. Review of fluorescence suppression techniques in Raman spectroscopy. *Applied Spectroscopy Reviews* **2015**, *50*, 387–406. doi:10.1080/05704928.2014.999936.
23. Weakley, A.T.; Griffiths, P.R.; Aston, D.E. Automatic baseline subtraction of vibrational spectra using minima identification and discrimination via adaptive, least-squares thresholding. *Applied Spectroscopy* **2012**, *66*, 519–529. doi:10.1366/110-06526.
24. Kloz, M.; Van Grondelle, R.; Kennis, J.T. Wavelength-modulated femtosecond stimulated raman spectroscopy - Approach towards automatic data processing. *Physical Chemistry Chemical Physics* **2011**, *13*, 18123–18133. doi:10.1039/c1cp21650c.
25. Cadusch, P.J.; Hlaing, M.M.; Wade, S.A.; McArthur, S.L.; Stoddart, P.R. Improved methods for fluorescence background subtraction from Raman spectra. *Journal of Raman Spectroscopy* **2013**, *44*, 1587–1595. doi:10.1002/jrs.4371.
26. Cooper, J.B.; Abdelkader, M.; Wise, K.L. Sequentially shifted excitation Raman spectroscopy: Novel algorithm and instrumentation for fluorescence-free Raman spectroscopy in spectral space. *Applied Spectroscopy* **2013**, *67*, 973–984. doi:10.1366/12-06852.
27. Ishikawa, S.T.; Gulick, V.C. An automated mineral classifier using Raman spectra. *Computers and Geosciences* **2013**, *54*, 259–268. doi:10.1016/j.cageo.2013.01.011.
28. Hermosilla Lopez-Reyes, G.; Catala, A.; Sanz, A.; Llanos, D. R.; Rull, F.; I. Raman spectra processing algorithms and database for RLS-ExoMars. *European Planetary Science Congress EPSC* **2012**, *7*, 1–2.
29. Jahoda, P.; Drozdovskiy, I.; Sauro, F.; Turchi, L.; Payler, S.; Bessone, L. Machine Learning for recognition of minerals from multispectral data **2020**. *1*, 1–11, [2005.14324].
30. De Lucia, F.C.; Harmon, R.S.; McNesby, K.L.; Winkel, R.J.; Mizolek, A.W. Laser-induced breakdown spectroscopy analysis of energetic materials. *Applied Optics* **2003**, *42*, 6148. doi:10.1364/ao.42.006148.
31. Rehse, S.J.; Salimnia, H.; Mizolek, A.W. Laser-induced breakdown spectroscopy (LIBS): An overview of recent progress and future potential for biomedical applications. *Journal of Medical Engineering and Technology* **2012**, *36*, 77–89. doi:10.3109/03091902.2011.645946.
32. Liu, J.; Osadchy, M.; Ashton, L.; Foster, M.; Solomon, C.J.; Gibson, S.J. Deep convolutional neural networks for Raman spectrum recognition: A unified solution. *Analyst* **2017**, *142*, 4067–4074, [1708.09022]. doi:10.1039/c7an01371j.
33. Acquarelli, J.; van Laarhoven, T.; Gerretzen, J.; Tran, T.N.; Buydens, L.M.; Marchiori, E. Convolutional neural networks for vibrational spectroscopic data analysis. *Analytica Chimica Acta* **2017**, *954*, 22–31. doi:10.1016/j.aca.2016.12.010.
34. Liang, J.; Mu, T. Recognition of big data mixed Raman spectra based on deep learning with smartphone as Raman analyzer **2020**. pp. 1–5. doi:10.1002/elps.201900302.
35. Atkinson, A.M.; Atkinson, A.M. ODU Digital Commons Demonstration of Visible and Near Infrared Raman Spectrometers and Improved Matched Filter Model for Analysis of Combined Raman Signals by. PhD thesis, Old Dominion University, 2019. doi:10.25777/e77s-3c37.
36. Thompson, D.R.; Castanfo, R. Performance Comparison of Rock Detection Algorithms for Autonomous Planetary Geology **2007**.
37. Thompson, D.R.; Allwood, A.; Bekker, D.; Cabrol, N.A.; Estlin, T.; Fuchs, T.; Kiri, L.; View, M. TextureCam: Autonomous Image Analysis for Astrobiology Survey **2012**. pp. 4–5.
38. Francis, R.; McIsaac, K.; Thompson, D.D.R.; Osinski, G.R.G. Autonomous Rock Outcrop Segmentation as a Tool for Science and Exploration Tasks in Surface Operations. *13th International Conference on Space Operations, SpaceOps 2014* **2014**, pp. 1–12. doi:10.2514/6.2014-1798.
39. Kissi, J.; Francis, R.; Pilles, E.; Osinski, G. Application Of Autonomous Science Targeting Using Texturecam For The Canmars 2016 Analogue Mission. *Lunar and Planetary Science Conference* **2017**, *48*.
40. Doerksen, K. Scholarship @ Western Machine Learning Classification of Interplanetary Coronal Mass Ejections Using Satellite Accelerometers. PhD thesis, 2019.
41. Bekker, D.L.; Thompson, D.R.; Abbey, W.J.; Cabrol, N.A.; Francis, R.; Manatt, K.S.; Ortega, K.F.; Wagstaff, K.L. Field Demonstration of an Instrument Performing Automatic Classification of Geologic Surfaces. *Astrobiology* **2014**, *14*, 486–501. doi:10.1089/ast.2014.1172.
42. Marion, C.L.; Osinski, G.R.; Bourassa, M.; Caudill, C.M.; Cloutis, E.; Christoffersen, P.; Hill, P.; Morse, Z.R.; Newman, J.; Pilles, E.A.; Simpson, S.L.; Tornabene, L.L.; Xie, T. OVERVIEW OF THE CANMOON LUNAR SAMPLE RETURN ANALOGUE MISSION. *51st Lunar and Planetary Science Conference* **2020**, pp. 2019–2020.

43. Pascual, A.; Kissi, J.; Osinski, G.R.; McIsaac, K. PIXEL-WISE CLASSIFICATION AND AUTONOMOUS IMAGE ANALYSIS IN A REAL-TIME ROVER OPERATIONS SCENARIO: LESSONS LEARNED FROM THE CANMOON ANALOGUE MISSION. *51st Lunar and Planetary Science Conference 2020*, pp. 51–52.
44. Shah, J.; Andres, C.; Pilles, E.A.; Roberts, A.; Rodríguez Sanchez-Vahamonde, C, D.; Vlachopoulos, O.; Yingling, W.; Marion, C.L.; Osinski, G.R.; Cloutis, E.; Morse, Z.R.; Newman, J.; Caudill, C.M. CANMOON ANALOGUE MISSION OPERATIONS: GIS, LOCALIZATION, AND ROVER TRAVERSE PLANNING AND MONITORING. *51st Lunar and Planetary Science Conference 2020*, 21, 1–9.
45. Caudill, C.M.; Osinski, G.R.; Pilles, E.; Sapers, H.M.; Pontefract, A.J.; Francis, R.; Duff, S.; Laughton, J.; O'Callaghan, J.; Sopoco, R.; Tolometti, G.; Tuite, M.; Williford, K.H.; Xie, T. Field and laboratory validation of remote rover operations Science Team findings: The CanMars Mars Sample Return analogue mission. *Planetary and Space Science* **2019**, *176*, 104682. doi:10.1016/j.pss.2019.06.006.
46. Bi, Y.; Zhang, Y.; Yan, J.; Wu, Z.; Li, Y. Classification and discrimination of minerals using laser induced breakdown spectroscopy and Raman spectroscopy. *Plasma Science and Technology* **2015**, *17*, 923–927. doi:10.1088/1009-0630/17/11/06.
47. Khan, S.; Ullah, R.; Khan, A.; Wahab, N.; Bilal, M.; Ahmed, M. Analysis of dengue infection based on Raman spectroscopy and support vector machine (SVM). *Biomedical Optics Express* **2016**, *7*, 2249. doi:10.1364/boe.7.002249.
48. Shan, M.; Cheng, Q.; Zhong, Z.; Liu, B.; Zhang, Y. Deep-learning-enhanced ice thickness measurement using Raman scattering. *Optics Express* **2020**, *28*, 48–56.
49. Carey, C.; Boucher, T.; Mahadevan, S.; Bartholomew, P.; Dyar, M.D. Machine learning tools for mineral recognition and classification from Raman spectroscopy. *Journal of Raman Spectroscopy* **2015**, *46*, 894–903. doi:10.1002/jrs.4757.
50. Yap, B.W.; Rani, K.A.; Abd Rahman, H.A.; Fong, S.; Khairudin, Z.; Abdulla, N.N. An application of oversampling, undersampling, bagging and boosting in handling imbalanced datasets. *Lecture Notes in Electrical Engineering* **2014**, *285 LNEE*, 13–22. doi:10.1007/978-981-4585-18-7_2.
51. Ganganwar, V. An overview of classification algorithms for imbalanced datasets. *International Journal of Emerging Technology and Advanced Engineering* **2012**, *2*, 42–47.
52. Kotsiantis, S.; Kanellopoulos, D.; Pintelas, P. Handling imbalanced datasets : A review. *Science* **2006**, *30*, 25–36, [arXiv:10.1007/978-0-387-09823-4-45].
53. He, K. Deep Residual Learning for Image Recognition. [arXiv:1512.03385v1].