UNIVERSITY OF TARTU

Institute of Computer Science

Computer Science Curriculum

Enriko Kroon

Chemical structure elucidation from nuclear magnetic resonance spectra using CAM methods with neural networks.

Bachelor's Thesis (9 ECTS)

Supervisor(s):
Stefan Hermann Kuhn, PhD

Tartu 2025

Chemical structure elucidation from nuclear magnetic resonance spectra using CAM methods with neural networks.

Abstract:

The given thesis investigates the application of visual explainability methods, specifically Grad-CAM and Grad-CAM++, for the identification of chemical substructures in two-dimensional NMR spectra using convolutional neural networks. A desktop application was developed to integrate these algorithms and analyse spectra from HMBC and HSQC experiments. The analysis revealed that chemical mixtures containing additional spectral components led to inconsistent and unreliable heatmaps. The study concludes that Grad-CAM and Grad-CAM++ combined with simple neural network architectures can highlight the pure compound spectra in most cases with varying accuracy, but are insufficient for reliably identifying fatty acid, indole, or steroid substructures in complex spectral mixtures.

Keywords: Chemoinformatics, Explainable AI, 2D NMR spectra, Grad-CAM, Grad-CAM++
CERCS: P175, Informatics, systems theory

Keemiliste struktuuride selgitamine tuumamagnetresonantsspektrites närvivõrkude ja CAM meetoditega.

Lühikokkuvõte:

Käesolevas lõputöös uuriti visuaalsete närvivõrkude seletatavusmeetodite, eelkõige Grad-CAM ja Grad-CAM++ algoritmide, rakendatavust keemiliste alamstruktuuride tuvastamisel kahedimensionaalsetes tuumamagnetresonantsspektrites, kasutades konvolutsioonilisi närvivõrke. Töö käigus arendati rakendus, mis võimaldab neid algoritme kasutada HMBC ja HSQC spektrite analüüsimiseks eelnevalt treenitud närvivõrkudel. Analüüsi tulemused näitasid, et puhaste keemiliste elementide spektrite puhul on enamik juhtudel võimalik välja tuua alamstruktuuri asukoht spektripildil varieeruva täpsusega. Spektrite puhul, mis sisaldasid täiendavaid komponente lisaks närvivõrgu poolt klassifitseeritavatele elementidele, oli tulemuste usaldusväärsus ja täpsus varieeruv. Töös järeldati, et Grad-CAM ja Grad-CAM++ algoritmid koos lihtsakoeliste närvivõrguarhitektuuridega ei ole piisavad rasvhapete, indoolide ega steroidide alamstruktuuride usaldusväärseks tuvastamiseks keerulistes spektrisegudes.

Võtmesõnad: Kemoinformaatika, Seletatavus tehisnärvivõrkudes, kahedimensionaalne tuumamagentresonantsspekter, Grad-CAM, Grad-CAM++

CERCS: P175, Informaatika, süsteemiteooria

Table of Contents

1. Definitions

NMR spectroscopy - Nuclear magnetic resonance spectroscopy

HMBC - Heteronuclear Multiple Bond Correlation

HSQC - Heteronuclear Single Quantum Coherence

CAM - Class activation map

2. Introduction

Chemistry is a discipline that demands the utmost accuracy as the slightest change in the process of synthesizing chemical compounds can have a great impact on the outcome of the chemical reaction. Such minor differences can have a major impact on the usability of the substance in pharmaceuticals, material science, academia and so on. Therefore it is crucial for people working in the chemical industry to have reliable methods for validating the presence of necessary substructures and molecules in chemical products.

One effective method for determining molecular structure is nuclear magnetic resonance (NMR) spectroscopy. NMR is a widely used non-destructive analytical technique in chemical research that helps to identify the structures of various compounds. It is based on the magnetic properties of atomic nuclei, which align when exposed to an external magnetic field and interact with radio frequency signals. By measuring the energy absorption of these nuclei, resonance signals can be obtained that generate an NMR spectrum. Depending on the specific NMR technique applied, different structural information can be extracted, ultimately aiding in the identification of a compound.

Images obtained through NMR spectroscopy can be used hand in hand with different specialized software to determine the exact content of a sample. The most common method that is used in the industry are lookup-tables based software like ACD / Labs Spectrus Processor [1]. By such software the NMR spectra is compared against other spectra samples in the database, and the content of the sample is determined through algorithmic comparison techniques. Although it is effective by pure substances, the database reliant software can struggle with identifying compounds from mixtures that include multiple chemical elements.

In addition to comparison based software, the use of neural networks is also explored by chemical structure elucidation. Neural networks can be used for classification tasks in hand with NMR spectra. These methods can be effective but lack explainability, as the spectra generated by chemical mixtures can be fairly difficult to analyze. The aim of the given bachelor thesis is to explore explainability methods by neural networks trained on chemical structure

1 ACD / Labs Spectrus Processor, https://www.acdlabs.com/products/spectrus-platform/spectrus-processor/, used 2025.

identification tasks to make the decision process of the network more interpretable and provide additional information about the chemical substructure and its location in the spectra.

3. NMR Spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy is a non destructive analysis method used to study the molecular structure of chemical compounds. It uses the magnetic properties of nuclei, which when placed in a magnetic field, resonate at characteristic frequencies depending on their chemical environment. Such effects are known as chemical shifts. The measured chemical shifts facilitate the analysis of the underlying chemical structure of the given element [1].

3.1 HSQC (Heteronuclear Single Quantum Coherence)

HSQC is a 2D NMR experiment that correlates hydrogen isotope nuclei with directly bonded heteronuclei, typically nitrogen or carbon isotopes, via one-bond couplings. Each peak in an HSQC spectrum corresponds to a specific 1H-X pair, where X is either 15N (Nitrogen-15) or 13C (Carbon-13). The HSQC experiment is especially valuable in biomolecular NMR imaging [2].

3.2 HMBC (Heteronuclear Multiple Bond Correlation)

HMBC is a 2D NMR experiment that detects correlations between hydrogen isotopes (1H) and heteronuclei, typically nitrogen (15N) or carbon (13C) isotopes, over two or more chemical bonds [2].

4. Previous research

Traditional NMR analysis relies on software tools that utilize databases of pre-defined spectral measurements to identify chemical elements. However advancements in machine learning have introduced new possibilities, particularly through the use of convolutional neural networks (CNNs) to analyze 2D NMR spectra directly as images. Kuhn et al. explored this approach and demonstrated that CNNs can successfully identify chemical substructures, like indole, fatty acids and steroids, from two-dimensional NMR spectra. Their research used spectral data from both Heteronuclear Multiple Bond Correlation (HMBC) and Heteronuclear Single Quantum Coherence (HSQC) experiments to train neural networks for substructure classification found in NMR spectra [3] .

To assess the effectiveness of this methodology, three separate CNN models were trained: one utilizing only HMBC spectra, another based only on HSQC spectra, and a third model incorporating both HMBC and HSQC spectra. The performance of these networks was evaluated through ten-fold cross-validation, which demonstrated that the model trained solely on HSQC spectra achieved an accuracy of 82.56 percent, whereas the HMBC-based model exhibited a higher accuracy of 91.86 percent. The combined model, which integrated both HSQC and HMBC spectra, achieved an accuracy of 90.84 percent. These findings indicate that HMBC spectra provide greater classification accuracy compared to HSQC alone, while the combination of both spectra gives roughly the same accuracy as using HMBC spectra alone [3].

Beyond classification of pure compounds, the study also investigated the ability of the trained networks to identify substructures in artificially generated mixtures. In this context, the HSQC-based model correctly classified four out of fifteen mixtures, the HMBC-based model successfully classified thirteen out of fifteen, and the combined model identified ten out of fifteen. The results suggest that HMBC spectra contain richer structural information for substructure classification than HSQC spectra alone. However, the study also found that certain substructures, particularly steroids, were more challenging to classify correctly, with some errors demonstrating consistent misclassification patterns.

5. Used technologies

In the following chapter a brief overview of the technologies used in the given thesis is given.

Python

Python[2] is a general-purpose interpreted programming language that features a dynamic type system and supports object-oriented programming. For the implementation of the practical part of the thesis, Python version 3.11.2 is used.

TensorFlow

TensorFlow[3] is a machine learning framework that can be used to build various types of neural networks. The framework also includes capabilities for parallel data processing and supports both multi-process training and GPU-based training. The TensorFlow API is available as a library for the Python programming language.

Tkinter

Tkinter[4] is a standard graphical user interface (GUI) library for the Python programming language. It provides tools for creating desktop applications with elements such as windows, buttons, menus, and input fields. Tkinter is included by default in Python distributions, making it convenient for developing user interfaces without requiring external libraries.

NumPy

NumPy[5] is an open source library that provides an API for performing numerical operations in Python. Due to its high performance it is widely used in scientific computing and when working with neural networks and machine learning.

2 Python, https://docs.python.org/3/faq/general.html
3 Tensorflow, https://www.tensorflow.org/about/bib
4 Tkinter, https://docs.python.org/3.11/library/tkinter.html
5 NumPy, https://numpy.org/about/

Pillow

Pillow[6] is an open source image processing library in Python. It supports multiple file formats and integrates well with other libraries that perform numerical operations on images like NumPy.

6 Pillow, https://pypi.org/project/pillow/

6. Explainability in neural networks

When neural networks produce an outcome of an analysis task, then the decision making process remains fairly abstract for the user to grasp - the prediction outcome may be correct, but were the features used for the decision making process correct or picked randomly? To answer that question there is a need for analysis methods that would help to explain the outcome of the prediction. One of the more common methods of explainability used with image processing neural networks are class activation maps (CAM) and their derivatives.

In the following chapter an overview of different class activation map methods will be given, that were considered for implementation in the scope of the given thesis. Grad-CAM and Grad-CAM++ were the methods chosen for implementation for the analysis software.

6.1 CAM

Class Activation Maps (CAM), originally introduced by Zhou et al., are used to visualize the regions in an input image that a convolutional neural network (CNN) uses when predicting a specific class [4].

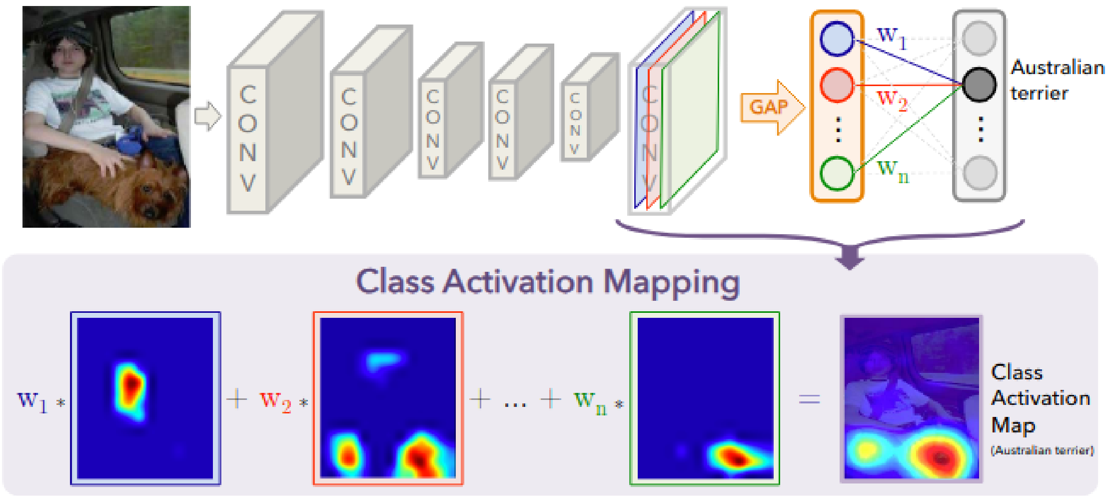

Figure 1. Class Activation Mapping: the predicted class score is mapped back to the previous convolutional layer to generate the class activation maps [4].

The method relies on information from the last convolutional layer of the network, just before the final classification layer, providing a high-level overview of the spatial features relevant to the prediction. In the original CAM method, the CNN architecture must include a Global Average Pooling (GAP) layer before the output layer. The GAP layer takes each feature map from the last convolutional layer and computes its spatial average, resulting in a single value per feature map (Figure 1). These values are then passed to a fully connected layer, where each class has a set of learned weights that indicate the importance of each feature map for that class. To generate the class activation map, these weights are multiplied by the corresponding original feature maps, and the weighted feature maps are summed to produce a heatmap. This heatmap highlights the regions of the image that were most influential in predicting the specific class [4].

6.2 Grad-CAM

Gradient-weighted class activation mapping (Grad-CAM) is a visual explainability technique proposed by Selvaraju et al., that uses weighted gradients of any target class to produce heat maps of the areas that the neural network deems important for classifying the given class. Unlike CAM, Grad-CAM does not require the presence of a global average pooling layer before the classification output, thus making it applicable to a wider variety of neural network architectures [5].

By Grad-CAM the importance of the feature maps are determined by calculating the gradient in respect to the majority class (or any target class of interest) of the feature maps. Then the given gradients are global-average-pooled over the width and height dimensions to get the importance of the neurons. The importance scores are then used as the weight of the corresponding feature maps. By combining these weighted maps, a single localized map is created that highlights the areas of the image that had the most significant impact on the network's decision. After that the result is passed through ReLu activation to obtain only the areas that had a positive effect on the network's decision making process thus creating a heat map [5].

6.3 Grad-CAM++

Gradient-weighted class activation mapping plus plus (Grad-CAM++) is a derivative and extension of the Grad-CAM algorithm proposed by Chattopadhay et al. [6]. Grad-CAM ++ tries to improve on the accuracy of the heatmaps by multiple object instances or small target regions. The main difference between Grad-CAM and Grad-CAM++ is the order of gradients used for determining the importance of the feature maps. In addition to first order gradients like by Grad-CAM, Grad-CAM++ also utilizes second and third order gradients of the feature map by the importance calculation. This improved weighting scheme ensures that feature maps contributing to smaller or multiple object regions are not underrepresented in the final saliency map. As a result, Grad-CAM++ produces sharper and more accurate visual explanations of CNN predictions, even in complex visual scenarios [6].

7. Methodology

The scope of the given bachelor’s thesis is to create a visual explainability workflow for the neural networks trained in the aforementioned study where chemical substructure classification models were implemented [3]. The aim was to see whether the individual substructure i.e. the majority class predicted by the neural network could be highlighted from the HMBC and HSQC spectra given as the prediction input, thus improving the usability of such neural networks by exact structure elucidation. In addition to producing an explainability method for such neural networks, generating class activation maps would also give insight by validating the networks function by finding out if the correct features are being studied by training.

7.1 Chosen CAM methods

For the purposes of this thesis, two class activation mapping algorithms were selected: Grad-CAM and Grad-CAM++. The primary factor influencing this choice was the architectural design of the neural networks employed as the foundation of this work. Specifically, the conventional CAM method requires the presence of a global average pooling (GAP) layer prior to the final output layer - a component that is absent in the models developed by Kuht et al. in their study. Consequently, CAM was deemed unsuitable for application within the context of these network architectures. Grad-CAM and Grad-CAM++ were chosen, as they do not have strict requirements for the underlying neural network’s architecture and provide a general approach to the neural network explainability problem.

7.2 Application structure

To facilitate the Grad-CAM and Grad-CAM++ algorithms it was necessary to create an application for performing the analysis on the spectra. Therefore a simple desktop application was created with a graphical user interface for selecting the spectra and generating heatmaps for them.

The framework of choice for the application created in the scope of the thesis was Tkinter, as it is included in the Python standard library. As the complexity of the application architecture is fairly minimal, it made sense to use something lightweight that already has a seamless integration with Python. Tkinter has a predefined set of utilities that come with the standard library that are called widgets in the context of the library. Widgets can be seen as a set of components or building blocks for the application's graphical user interface. One can also create custom widgets that serve as custom components for the application where the custom widget is just an extension of a predefined base class in Python.

For the application GUI two custom widgets were created - ControlPanel and ImageLoader. The ImageLoader widget handles the operations related to loading and displaying the spectra. The class handles loading, updating and saving the image. The ControlPanel widget handles all the user actions related to performing the analysis on the image. It requires an ImageLoader class instance to be injected into the class constructor as it calls the methods associated with updating the image in the ImageLoader widget. The ControlPanel class also provides the ability to select preferences for the user. The user can choose between different pre-trained neural networks, CAM methods (Grad-CAM and Grad-CAM++) and also the fully connected layer of the neural network which will be used for the CAM analysis. The ControlPanel also handles the communication with the AnalysisService class which handles the image preprocessing tasks.

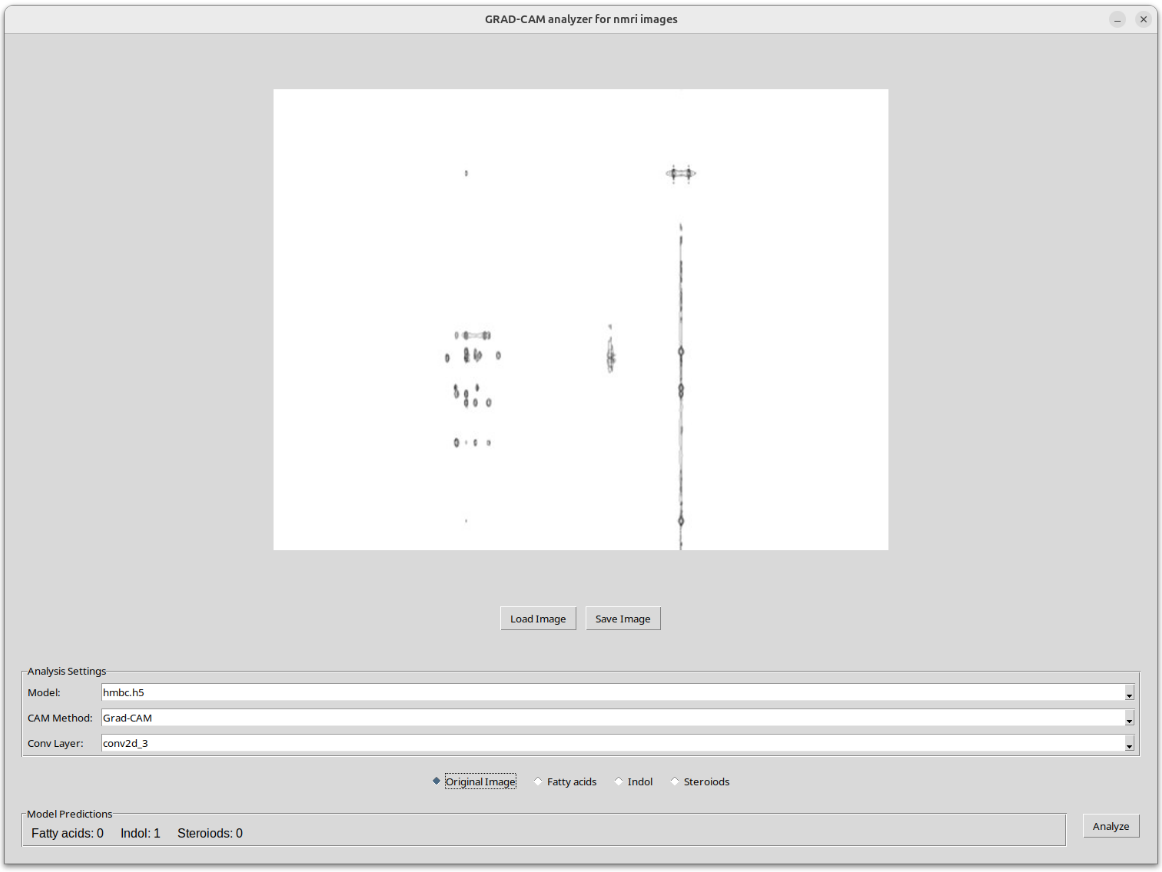

Figure 2. Screenshot of the user interface of the application.

The AnalysisService class serves as the backbone of the applications. It handles the necessary function calls related to performing the CAM analysis on the image. The CAM methods can be defined as separate functions and later imported to the AnalysisService class to be used by the user.

The neural network models and datasets utilized in this project were sourced from the study by Kuhn et al.[7] The project's repository was cloned from GitHub, and after addressing issues related to deprecated dependencies, the models for HMBC and HSQC classification were successfully trained and exported.

The final application created in the scope of the given bachelor's thesis can be found on GitHub under the name nmri-grad-cam[8].

7 substructuresnm, https://github.com/stefhk3/substructuresnmr/tree/master

8 nmri-grad-cam, https://github.com/EnrikoK/nmri-grad-cam

8. Analysis of CAM methods

To evaluate the effectiveness of Class Activation Mapping (CAM) methods and neural networks in chemical structure elucidation, a series of heatmap images were generated using a previously developed analysis workflow. The evaluation focuses on comparing the heatmaps of pure substances with those of artificially created mixtures that include the original pure substances. These artificial mixtures are produced by overlaying spectra from the three target compound classes—steroids, fatty acids, and indoles—with spectra from compounds that do not belong to any of these subclasses. The spectra used were taken from the initial research paper by Kuhn et al. which belonged to the neural network training and testing image set [3] . Each artificial mixture will include spectral data from one of the target classes along with one, or two additional spectra from unrelated compounds. This method enables the assessment of the models' ability to accurately classify relevant substructures and identify the substructure region in the spectra using Grad-CAM and Grad-CAM++ algorithms. The procedure will be carried out separately for both HMBC and HSQC spectra.

8.1 Analysis of Grad-CAM

8.1.1 Pure substances

By evaluating pure substances, the neural network managed to correctly classify all of the substances by both HMBC and HSQC spectra (Figures 3, 4, 5, 6, 7, 8). The generated heat map images showed minor misalignment with the actual location of the spectra. This could be because of the upscaling of the heatmap when it is generated from the last dense layer of the neural network.

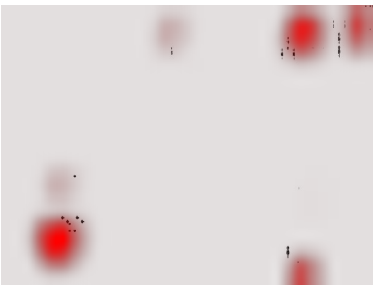

Figure 3. Grad-CAM heatmap of HMBC fatty acid spectra.

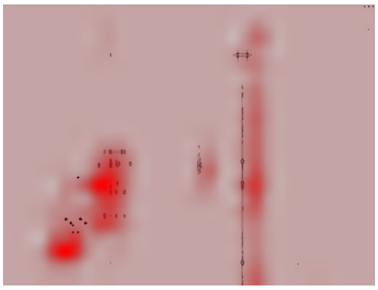

Figure 4. Grad-CAM heatmap of HMBC indole spectra.

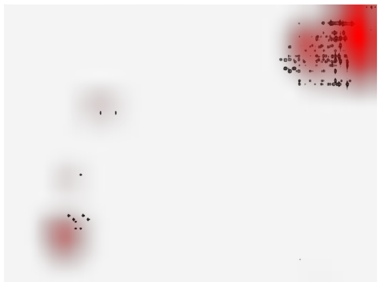

Figure 5. Grad-CAM heatmap of HMBC steroid spectra

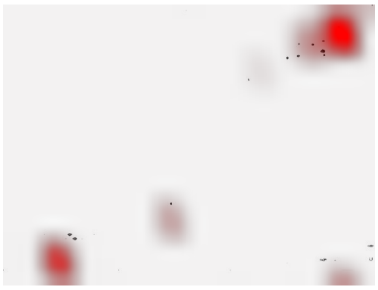

Figure 6. Grad-CAM heatmap of HSQC fatty acid spectra.

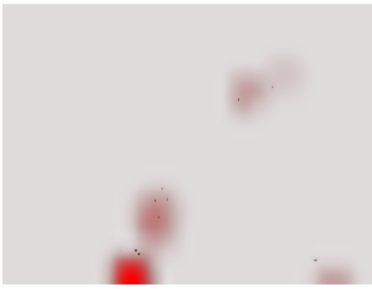

Figure 7. Grad-CAM heatmap of HSQC indole spectra.

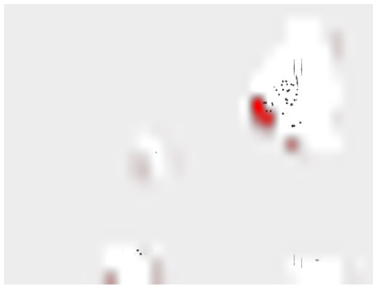

Figure 8. Grad-CAM heatmap of HSQC steroid spectra.

Although the heat maps generally highlight all major regions of interest, it is evident that certain spectral areas containing minor signals are given less emphasis in the network's decision-making process. In the case of the HSQC model, the Grad-CAM heatmap of the steroid spectra reveals that only a portion of the visible spectral cluster is considered relevant for the prediction, with a substantial amount of information left unaccounted for. Furthermore, for the HMBC spectra of an indole compound, it is noteworthy that the network's decision also appears to incorporate white regions—areas devoid of spectral information—suggesting a possible misinterpretation by the model.

8.1.2 Mixture with one additional compound

When utilizing artificial mixtures for analysis, it becomes evident that the network retains its ability to accurately classify the presence of the three target substructures. Upon generating the corresponding class-specific heatmaps, it is observed that the additional spectral information introduced by the mixtures exerts only a minimal influence on the network's decision-making process. The heatmaps reveal that regions containing spectral signals from non-target chemical compounds—those not belonging to the three designated classes—may exhibit slight activation. However, these areas are consistently less emphasized compared to regions associated with the pure spectra of the target compounds, indicating that the network in most cases focuses on relevant spectral features.

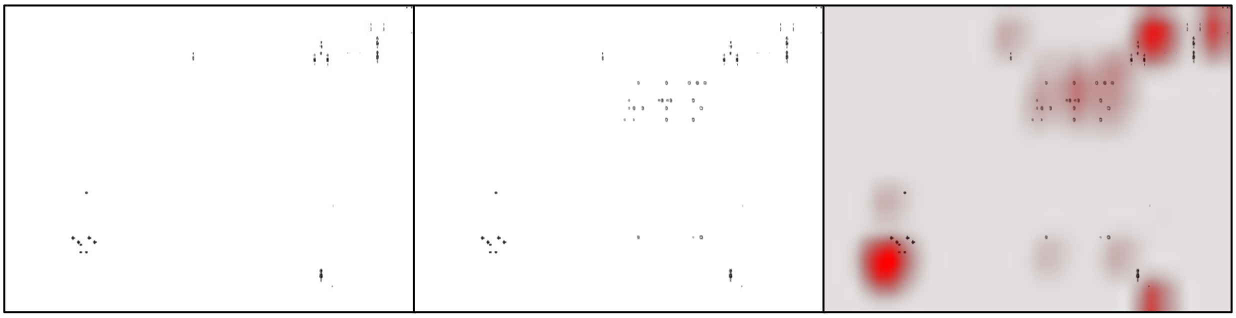

Figure 9. Fatty acid HMBC spectra, artificial mixture with bmse000123 (Trans 4 Hydroxy-L-proline ($C_5H_9NO_3$)) and Grad-CAM analysis of the mixture.

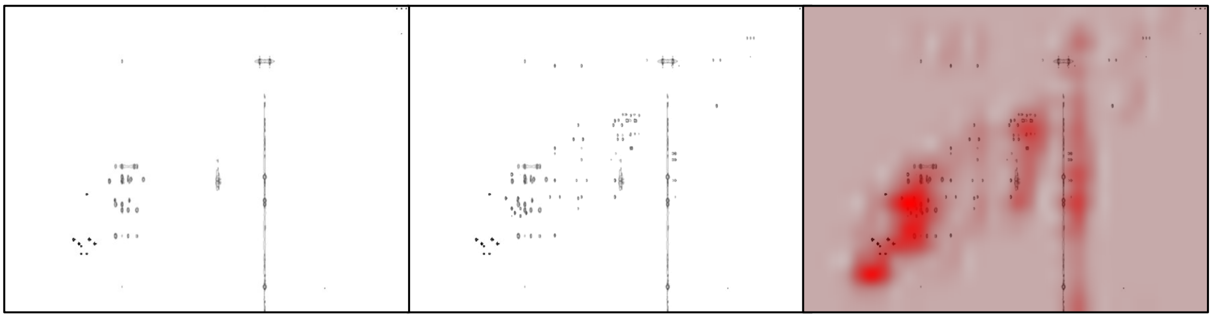

Figure 10. Indole HMBC spectra, artificial mixture with bmse000054 (NADH ($C_{21}H_{29}N_7O_{14}P_2$)) and Grad-CAM analysis of the mixture.

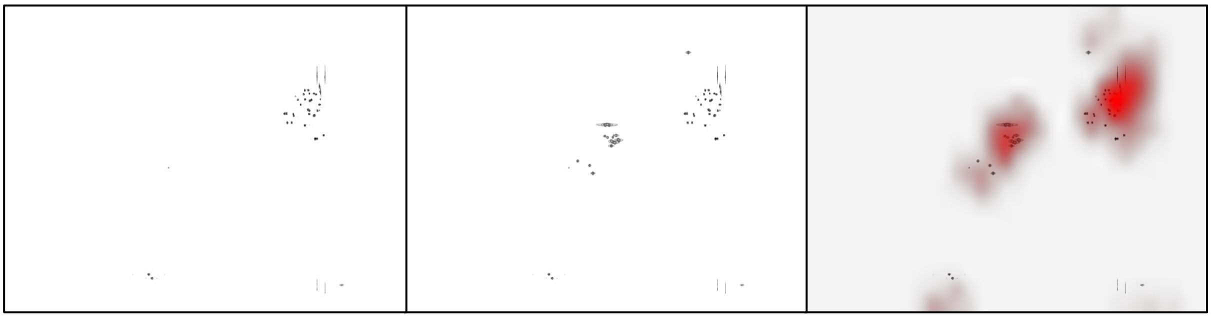

Figure 11. Steroid HSQC spectra, artificial mixture with bmse000138 (D-Cellobiose ($C12H22O11$)) and Grad-CAM analysis of the mixture.

As it can be seen in Figure 9 and Figure 11, the Grad-CAM algorithm also captures the additional spectral data that does not belong to the initial pure compound. By the HMBC fatty acid mixture with bmse000123 (Figure 9), the additional spectral information is only considered slightly important by the algorithm whereas by the HSQC steroid mixture (Figure 11) the additional spectral information from a non target class is considered almost equally important by the Grad-CAM algorithm and the neural network.

8.1.3 Mixture with two additional compounds

When there are more than one additional spectra included into the image then again the heatmaps vary in accuracy. The network is mostly able to classify correctly the existence of a substructure. One exception occurred when classifying the artificial mixture of a fatty acid with two additional compounds where the network misclassified it as a steroid substructure . There it can also be seen that the network focuses on the surrounding areas of the spectral information instead of the actual signal data (Figure 12).

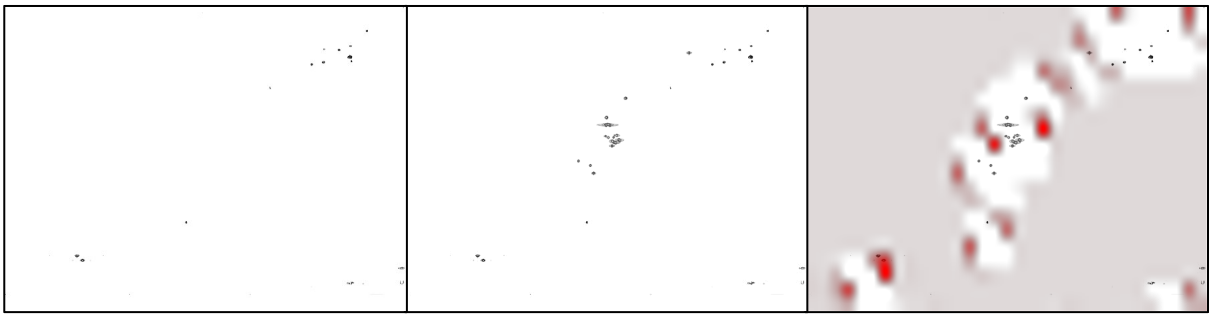

Figure 12. Fatty acid HSQC spectra, artificial mixture with bmse000078 (Creatine ($C_4H_9N_3O_2$)) and bmse000138 (D-Cellobiose ($C_{12}H_{22}O_{11}$)), Grad-CAM analysis of the artificial mixture. Misclassified as steroid.

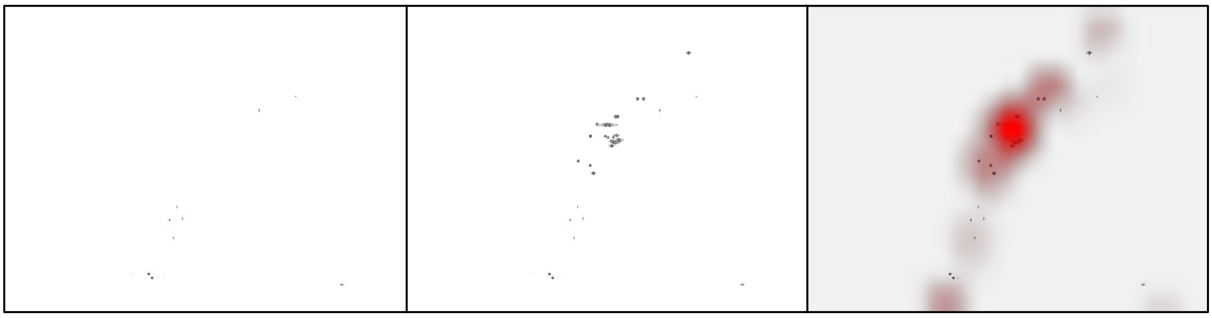

Figure 13. Indole HSQC spectra, artificial mixture with bmse000123 (Trans 4 Hydroxy-L-proline ($C_5H_9NO_3$)) and bmse000138 (D-Cellobiose ($C_{12}H_{22}O_{11}$)), Grad-CAM analysis of the artificial mixture.

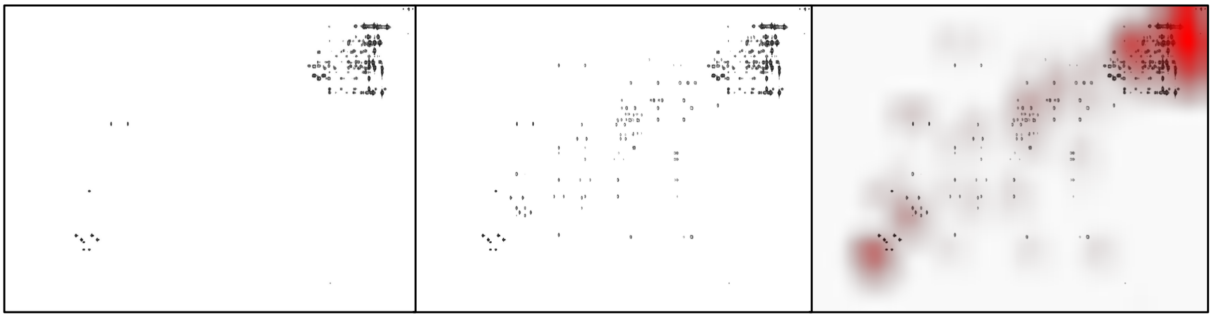

Figure 14. Steroid HMBC spectra, artificial mixture with bmse000054 (NADH ($C_{21}H_{29}N_7O_{14}P_2$)) and bmse000123 (Trans 4 Hydroxy-L-proline ($C_5H_9NO_3$)), Grad-CAM analysis of the artificial mixture.

It can also be seen that by mixtures with more than one additional spectra, the effectiveness of Grad-CAM can vary significantly. In Figure 13 the neural network along with Grad-CAM analysis also includes non relevant spectral information by the decision although it classified the substructure correctly. Also in Figure 14 the underlying steroid substructure was classified correctly from the HMBC spectra, but the generated heatmap is more accurate in a sense as the areas of the steroid spectra show greater activation.

8.2 Analysis of Grad-CAM++

8.2.1 Pure substances

By applying the Grad-CAM++ algorithm, the heat maps generated for the pure spectra cover almost all areas where signals are present in the spectral images (Figures 15, 16, 17, 18, 19, 20). While minor misalignments can still be observed, the clusters of spectral data generally produce heat maps that roughly align with the signal regions.

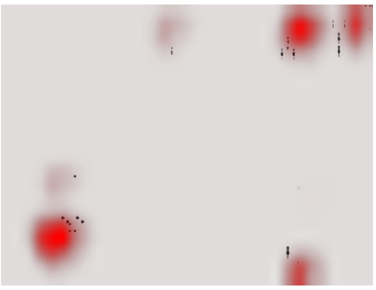

Figure 15. Grad-CAM++ heatmap of HMBC fatty acid spectra.

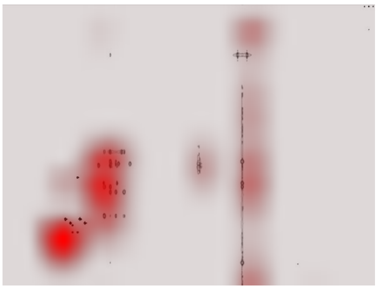

Figure 16. Grad-CAM++ heatmap of HMBC indole spectra.

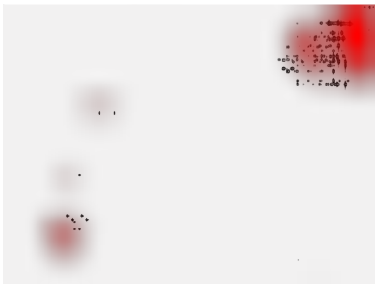

Figure 17. Grad-CAM++ heatmap of HMBC steroid spectra

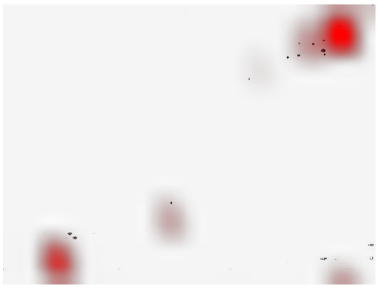

Figure 18. Grad-CAM++ heatmap of HSQC fatty acid spectra.

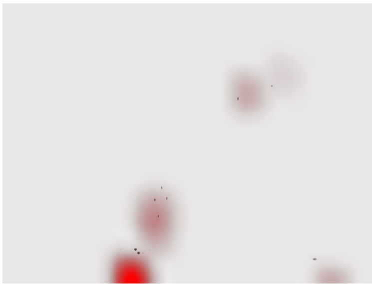

Figure 19. Grad-CAM++ heatmap of HSQC indole spectra.

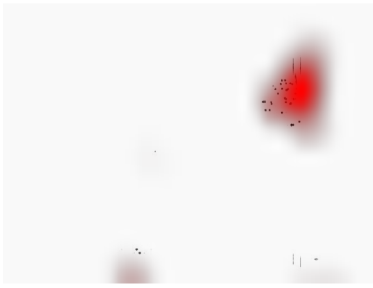

Figure 20. Grad-CAM++ heatmap of HSQC steroid spectra.

Smaller signal clusters mostly generate heatmaps slightly less intense compared to larger clusters, indicating that the algorithm and the neural network mostly weigh the importance of signals rather than simply their abundance or size of the information cluster.

8.2.2 Mixture with one additional compound

When analysing mixtures with one additional compound, the neural network is able to mostly classify the substructure correctly. Yet again the accuracy of the heat maps varies.

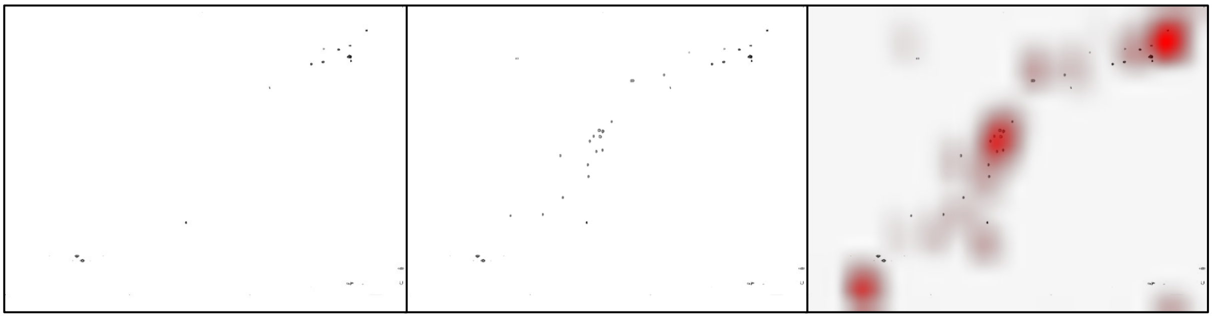

Figure 21. Fatty acid HSQC spectra, artificial mixture with bmse000054 (NADH ($C_{21}H_{29}N_7O_{14}P_2$)), Grad-CAM++ analysis of the artificial mixture. Misclassified as steroid.

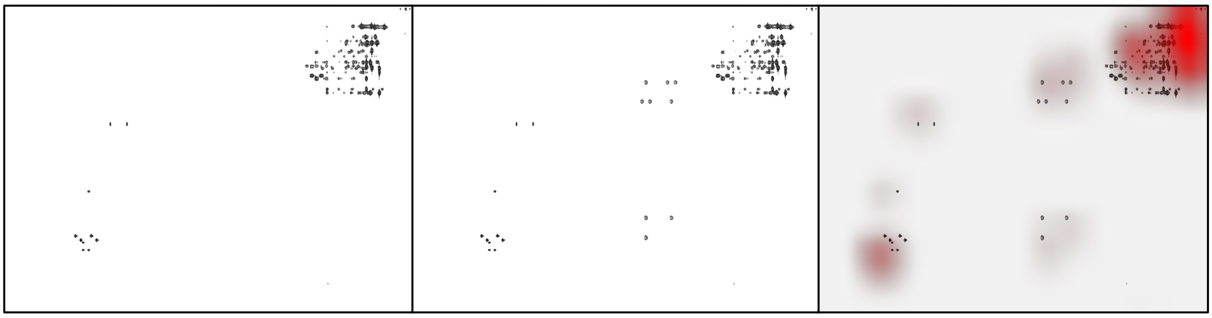

Figure 22. Steroid HMBC spectra, artificial mixture with bmse000078 (Creatine ($C_4H_9N_3O_2$)), Grad-CAM++ analysis of the artificial mixture.

In Figure 22 the neural network focuses on the larger cluster of information in the top right corner. Other spectral elements that do not belong to the given steroid substructure present also minor activation. In Figure 21 it is evident that the area in the middle of the spectral image is considered important for the neural network although it does not belong to the initial fatty acid substructure.

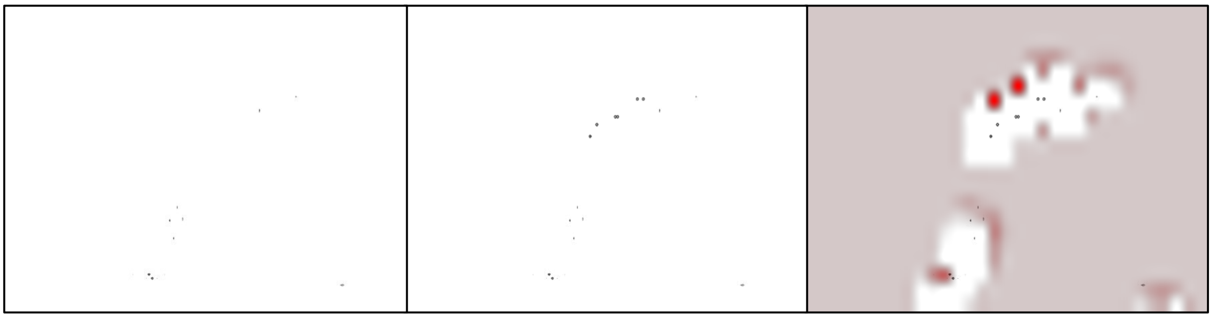

Figure 23. Indole HSQC spectra, artificial mixture with bmse000123 (Trans 4 Hydroxy-L-proline (C5H9NO3)), Grad-CAM++ analysis of the artificial mixture.

Also it can be seen that Figures 8, 12 and 23 present similar issues with the heat maps - the actual spectral information is considered minimally, but the surrounding white area is important for the network's decision making.

8.2.3 Mixture with two additional compounds

By two additional components in the artificial mixture the results do not differ much from the previously analysed heatmaps. The neural network is able to classify the substructure mostly correctly, but the accuracy of the alignment of the heatmaps with the actual substructure varies.

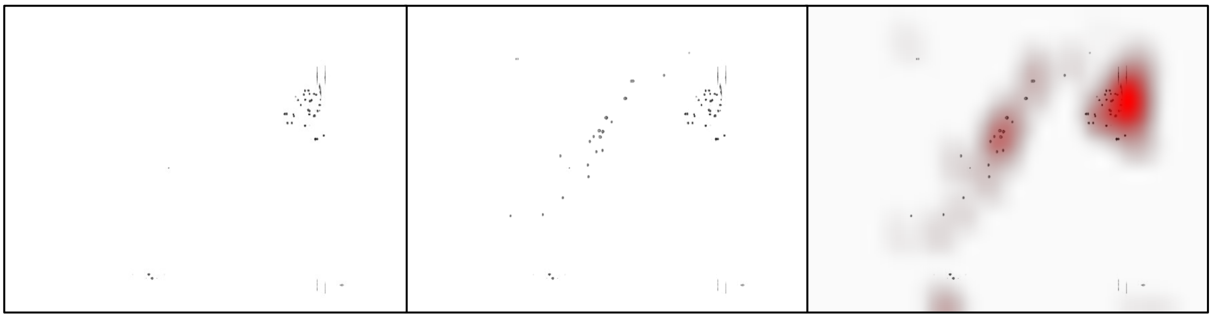

Figure 24. Steroid HSQC spectra, artificial mixture with bmse000054 (NADH (C21H29N7O14P2)) and bmse000078 (Creatine (C4H9N3O2)), Grad-CAM++ analysis of the artificial mixutre.

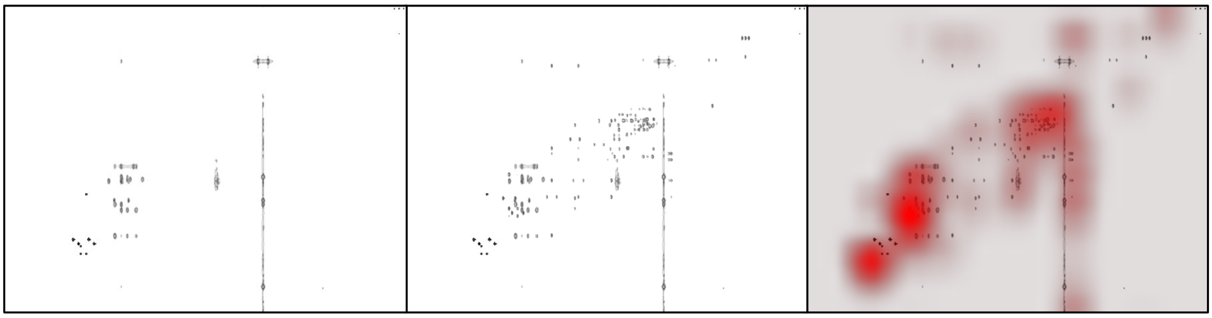

Figure 25. Indole HMBC spectra, artificial mixture with bmse000054 (NADH ($C_{21}H_{29}N_7O_{14}P_2$)) and bmse000138 (D-Cellobiose ($C_{12}H_{22}O_{11}$)), Grad-CAM++ analysis of the artificial mixture.

In Figure 24 the additional spectra that came with the artificial mixture have a minor effect on the neural networks decision making, the original steroid substructure gets highlighted but not that accurately. The same case is evident in Figure 25 – most of the original substructure spectra get highlighted by the heat map, but so do elements that get included with the artificial mixture.

9. Discussion

The main goal of the given thesis was to see whether the Grad-CAM and Grad-CAM++ algorithms could produce heat maps to locate specific substructures from spectra and artificial mixtures. The produced heat maps varied in accuracy by both Grad-CAM and Grad-CAM++. In addition to that fatty acid, indole nor steroid substructures could not be consistently identified from the artificial mixtures by both HMBC nor HSQC spectra.

By the analysis of the pure spectra, it is evident that the neural network is mostly able to consider most of the spectral information by the decision making process although with varying importance. It can be seen that the steroid HMBC spectra that has a larger grouping in the top right corner of the spectral image, produces consistent activation in the given area (Figures 5, 17). This remains so also by mixtures with one additional spectra that do not belong to the steroid class (Figure 22). The same seems to be the case with the HSQC steroid spectra where a larger grouping on spectral data can be seen in the top right corner of the spectral image (Figures 11, 20, 24) with the exception of the Grad-CAM heatmap of the pure HSQC steroid substructure where the larger grouping gets excluded from the heat map (Figure 8).

The heatmaps produced by the HSQC spectra tend to have more inconsistencies - the area of the heat map activation in some cases surrounds the cluster of spectral data instead of being laid on top of the cluster indicating its importance (Figures 8, 12, 23). The reason could be the lesser accuracy of the neural network trained purely on the HSQC data which was described in the aforementioned study [3].

Overall, for mixtures, the accuracy of the generated heat maps varied greatly for both Grad-CAM and Grad-CAM++. Images containing larger clusters of information within their pure spectra tended to produce heatmaps with higher activation concentrated in those areas. In contrast, smaller signal groupings in the spectral images generally resulted in heatmap regions with lower activation.

It was observed that the generated heatmaps could not be reliably used to identify the original substructures from the artificial mixtures. Additionally, spectral data incorporated from other compounds, that did not belong to the target classes, into the mixtures also produced activation in the neural networks to varying degrees, thus contributing to the heatmaps (Figures 9 ,10, 11, 12, 13, 21, 22, 23, 24, 25).

10. Conclusion

This thesis investigated the use of Grad-CAM and Grad-CAM++ algorithms for identifying chemical substructures in 2D NMR spectral data through simple neural network architectures. An analysis workflow with a supporting desktop application was developed to visualize the decision-making processes of neural networks trained on HMBC and HSQC spectra.

The experiments demonstrated that, although pure spectral data could mostly be visualized with varying accuracy on its own by the heatmaps. The generated heat maps for artificial mixtures containing additional spectra not originating from the three initial target classes varied significantly in accuracy. Neither Grad-CAM nor Grad-CAM++ consistently highlighted the correct substructures within mixtures, and non-target spectral signals often influenced the class activation maps.

These results suggest that the use of Grad-CAM and Grad-CAM++ with simple neural network architectures, similar to those described in the study by Kuhn et al., is not a feasible approach for identifying fatty acid, indole, nor steroid substructures in HMBC nor HSQC spectra when dealing with chemical mixtures containing additional spectral contributions from non target classes of the neural network. Future work could explore the use of more complex neural network architectures or other explainability methods to enhance interpretability.

References

[1] Nuclear magnetic resonance (NMR). In: Daintith J, editor. *A Dictionary of Chemistry*. 6th ed. Oxford: Oxford University Press, 2008, cited 2025; Available from: https://www.oxfordreference.com/display/10.1093/acref/9780199204632.001.0001/acref-9780199204632-e-2929

[2] Emwas AH, Roy R, McKay RT, Tenori L, Saccenti E, Gowda GN, Raftery D, Alahmari F, Jaremko L, Jaremko M, Wishart DS. NMR spectroscopy for metabolomics research. Metabolites, 2019, cited 2025; Available from: https://doi.org/10.3390/metabo9070123

[3] Kuhn S, Tumer E, Colreavy-Donnelly S, Borges RM. A Pilot Study For Fragment Identification Using 2D NMR and Deep Learning. Magn Reson Chem, 2021, cited 2025; Available from: https://doi.org/10.1002/mrc.5212

[4] Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning Deep Features for Discriminative Localization. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Computer Vision and Pattern Recognition (CVPR), 2016, cited 2025; Available from: https://doi.org/10.1109/CVPR.2016.319

[5] Selvaraju RR, Das A, Vedantam R, Cogswell M, Parikh D, Batra D. Grad-CAM: Why did you say that? 2016 Nov 22, cited 2025; Available from: https://doi.org/10.48550/arXiv.1611.07450

[6] Chattopadhay A, Sarkar A, Howlader P, Balasubramanian VN. Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Applications of Computer Vision (WACV), 2018 IEEE Winter Conference on, WACV, 2018, cited 2025. Available from: https://doi.org/10.1109/WACV.2018.00097

Appendices

GitHub repository for the created workflow, nmri-grad-cam https://github.com/EnrikoK/nmri-grad-cam

License

I, Enriko Kroon,

1. grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the digital archives of the University of Tartu until the expiry of the term of copyright, my thesis “Chemical structure elucidation from nuclear magnetic resonance spectra using CAM methods with neural networks.”, supervised by Stefan Hermann Kuhn.

2. grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright;

3. am aware of the fact that the author retains the rights specified in points 1 and 2;

4. confirm that granting the non-exclusive licence does not infringe other persons’ intellectual property rights or rights arising from the personal data protection legislation.

Enriko Kroon

11.05.2025