



Imam Khomeini International University
Vol. 9, No. 4, Winter 2024



نشریه مهندسی منابع معدنی
Journal of Mineral Resources Engineering
(JMRE)

Research Paper

Automatic Separation of Ore and Waste Using Images of Drill Cores and the Deep Neural Network U-Net

Farahmand M.¹, Fahiminia M.², Asghari O.^{3*}, Mirmohammadi M.S.⁴

1- M.Sc, Dept. of Mining, Faculty of Engineering, University of Tehran, Tehran, Iran

2- Ph.D, Dept. of Mining, Faculty of Engineering, University of Tehran, Tehran, Iran

3- Associate Professor, Dept. of Mining, Faculty of Engineering, University of Tehran, Tehran, Iran

4- Assistant Professor, Dept. of Mining, Faculty of Engineering, University of Tehran, Tehran, Iran

Received: 12 Jul. 2023

Accepted: 23 Apr. 2024

Abstract: One of the most crucial steps in ore exploration is the recognition of geological patterns and features. These features contain mineralogy, lithology, alteration, rock texture, etc. This stage has always been associated with many challenges. Among the challenges of this stage, we can mention the time-consuming and costly nature of this stage and the need for high expertise and human resources to recognize these patterns and features. In recent years, deep learning and machine learning have been adopted in earth sciences. In this research, by using the architecture of U-net, ore and waste were separated, and the grade pattern was identified using the core box images. For this purpose, iron minerals were segmented using binary image segmentation, trial-and-error methods were used to optimize the network, and finally, the model's accuracy for identifying ore was 91%. The IoU metric was utilized for further evaluation; this metric is a suitable criterion for the final evaluation of the image segmentation model, which has reached 75% in recognition of iron ores. For the final evaluation of the obtained model, the grade outputs of the model and the XRF analysis results of one core were compared. The network error was evaluated at 9%, which shows the good accuracy of the obtained model according to the real data.

Keywords: Core box, Image segmentation, RGB image, Deep learning, U-net network.

How to cite this article

Farahmand, M., Fahiminia, M., Asghari, O., and Mirmohammadi, M. S. (2024). "Automatic separation of ore and waste using images of drill cores and the deep neural network U-net". Journal of Mineral Resources Engineering, 9(4): 1-15.

DOI: [10.30479/JMRE.2024.19041.1652](https://doi.org/10.30479/JMRE.2024.19041.1652)

*Corresponding Author Email: o.asghari@ut.ac.ir

COPYRIGHTS



©2024 by the authors. Published by Imam Khomeini International University.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0/>)

INTRODUCTION

In mineral exploration, the drill core stands as the most reliable and important primary data, representing direct information from underground. Geological core logging is an essential part of mineral exploration, as it helps geologists and mining engineers determine the size, shape, and mineral composition of an ore body. This geological logging is often done visually and by a geologist. It is a time-consuming and labor-intensive task, and expert biases can affect it; thus, automated approaches are needed. In recent years, various research has been carried out in the field of automatic detection of minerals using color images, which in some traditional methods of machine learning [1,2] and in newer research deep learning methods based on convolutional neural networks (CNN) are utilized [3,4].

In this research, we used the U-net network [5], which is one of the CNN-based architectures, and image segmentation to recognize iron ore and waste in colored images of core boxes. Image segmentation is the process of partitioning a digital image into multiple segments to simplify or change the representation of the image. It is typically used to locate objects and boundaries in images. For the training phase, images of the core boxes were divided into suitable sizes, labeled by the geologist, and used as the input of the neural network. The outputs of the network were evaluated by accuracy and IoU metrics. Additionally, the output of the one core was compared to the XRF analysis results for validation.

METHODS

The automatic method used to separate ore and waste in this research is based on image segmentation. The first stage of the research is the collection of image data from the drilling core box, their analysis using the opinion of a geologist, and the results of chemical analysis. The images were divided into 256×256 patches and labeled. The generated labels indicate areas of iron ores for each pixel of the image. The image segmentation model in this research was U-net, which is a CNN-based network suitable for image segmentation tasks. U-net needs less data compared to other CNN models, and it can capture both coarse and fine feature information [5].

The U-net is a network with a symmetrical architecture and consists of four main parts, which include the contracting path, the expanding path, the bottleneck, and the output layer. Figure 1 shows an overview of the U-net along with its four main parts and different layers, where the vertical length of the layers represents the image dimensions and the horizontal length of the layers represents the number of feature maps.

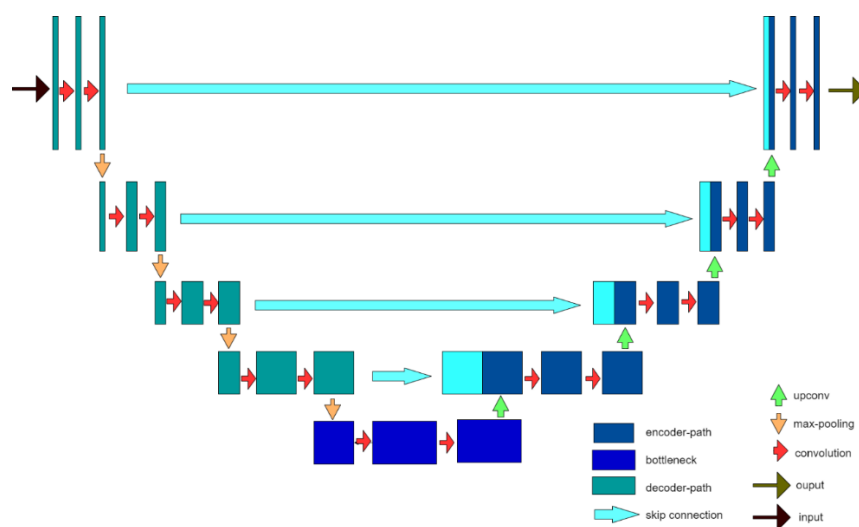


Figure 1. Architecture of U-net

The U-Net model consists of a contracting path (encoder) and an expansive path (decoder). The contracting path captures context and features from the input image while reducing its spatial resolution through convolutional and pooling layers, enabling the extraction of features at different scales. Meanwhile,

the expansive path reconstructs the segmented image to its original resolution using upsampling and convolutional layers. It refines details and combines information from different scales, facilitated by skip connections that preserve spatial information and merge low-and high-level features, resulting in precise pixel-wise segmentation. The bottleneck connects the encoder and decoder and has the most feature channels in it, and the output layer with one 1×1 convolution layer does the final classification or segmentation. The activation function of every layer is ReLU, but the final layer is sigmoid.

FINDINGS AND ARGUMENT

U-net neural network codes and image segmentation steps were done in the Python programming language, and no other external software was used. Also, the Tensorflow GPU environment version and the Python 3.9 version were used to build the U-net neural network. Figure 2 shows the loss and accuracy curves for training U-net. The curves show that the network is well-trained, while the error and accuracy curves are well-converged, and no overfitting has occurred.

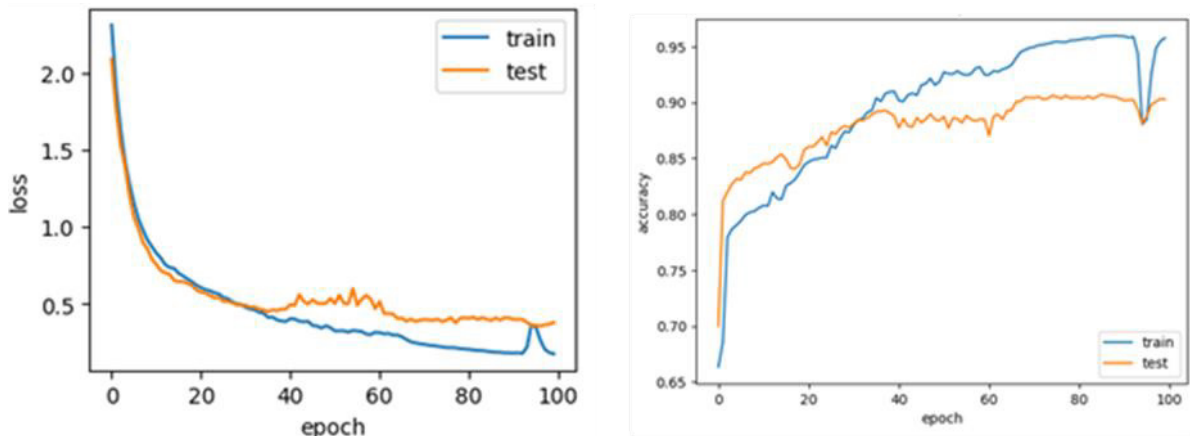


Figure 2. Loss and accuracy curve of U-net training

The output of the network is segmented images, where the defined segments represent iron mineralization. Figure 3 shows an instance of network prediction. To measure the performance of the network qualitatively, the labeled images, the original images, and the output images of the network were compared.

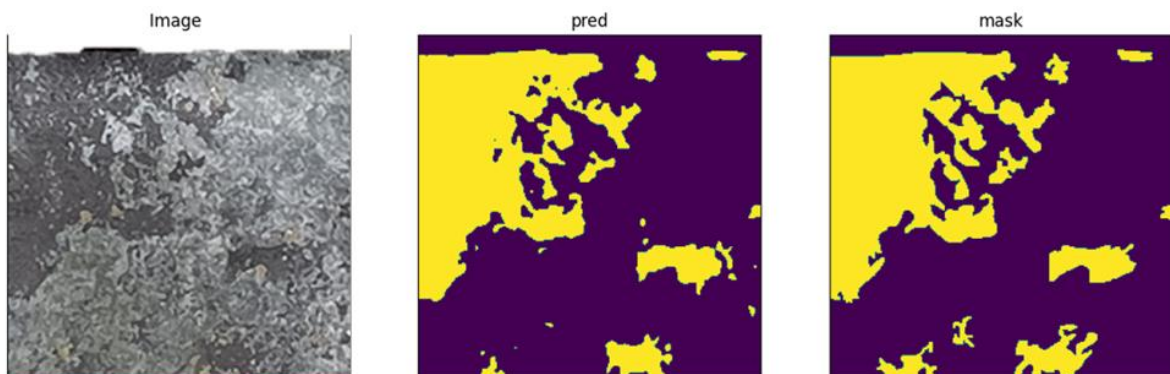


Figure 3. Ore and waste prediction of U-net with labeled mask and input image. The yellow color represents the ore, and the reset is waste

The final validation of the research was comparing the percentage of ores in prediction outputs with the XRF analysis data. Figure 4 shows an instance of this validation, where the 3rd row represents the percentage of yellow pixels (iron ore). The iron ore is magnetite in this core, and the average pixel percentage is 38.52%, resulting in 27.87% of Fe total. Compared to the 30.87% Fe total from the XRF analysis, the

automated method of ore detection had an acceptable performance.

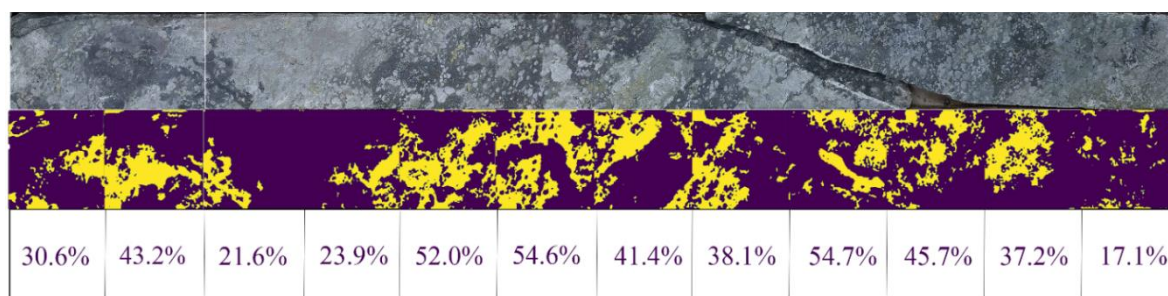


Figure 4. Visual and pixel percent validation

CONCLUSIONS

In this research, we developed an automated ore and waste detection approach for iron ore images. The method was based on a convolutional neural network, which is a deep learning algorithm. The result shows that this method can be utilized on different iron ore corebox images, and with an acceptable error, it can predict the boundaries of ore and waste. Although training deep learning networks demands a substantial volume of data, augmenting both the quantity and resolution of images can result in enhanced accuracy and reduced error rates. The weakness of the method is locations such as fractures that are not common in all images, and more images for training can solve this problem.

REFERENCES

- [1] Liu, J., Chen, W., Muller, M., Chalup, S., and Wheeler, C. (2019). "An automatic HyLoggerTM mineral mapping method using a machine-learning-based computer vision technique". *Australian Journal of Earth Sciences*, 66(7): 1063-1073. DOI: 10.1080/08120099.2019.1600167.
- [2] Desta, F., and Buxton, M. (2017). "The use of RGB Imaging and FTIR Sensors for Mineral mapping in the Reiche Zeche underground test mine, Freiberg". *Real Time Mining, Conference on Innovation on Raw Material Extraction*, 103-127.
- [3] Jin, C., Wang, K., Han, T., Lu, Y., Liu, A., and Liu, D. (2022). "Segmentation of ore and waste rocks in borehole images using the multi-module densely connected U-net". *Computers & Geosciences*, 159: 105018. DOI: 10.1016/j.cageo.2021.105018.
- [4] Alzubaidi, F., Mostaghimi, P., Swietojanski, P., Clark, S. R., and Armstrong, R. T. (2021). "Automated lithology classification from drill core images using convolutional neural networks". *Journal of Petroleum Science and Engineering*, 197: 107933. DOI: 10.1016/j.petrol.2020.107933.
- [5] Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: Navab, N., Hornegger, J., Wells, W., and Frangi, A. (Eds.), *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*. MICCAI 2015, Lecture Notes in Computer Science, Vol. 9351, Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-24574-4_28.



تشخیص اتوماتیک کانسنگ از باطله با استفاده از تصاویر مغزه‌ها و شبکه عمیق U-Net

مجتبی فرمند^۱، محمد فهیمی‌نیا^۲، امید اصغری^۳، میر صالح میرمحمدی^۴

- ۱- دانشجوی کارشناسی ارشد، دانشکده مهندسی معدن، دانشگاه تهران، تهران
- ۲- دانشجوی دکتری، دانشکده مهندسی معدن، دانشگاه تهران، تهران
- ۳- دانشیار، دانشکده مهندسی معدن، دانشگاه تهران، تهران
- ۴- استادیار، دانشکده مهندسی معدن، دانشگاه تهران، تهران

پذیرش: ۱۴۰۳/۰۲/۰۴

دریافت: ۱۴۰۲/۰۴/۲۱

چکیده

یکی از مراحل اولیه و مهم در اکتشاف مواد معدنی، تشخیص الگوها و ویژگی‌های زمین‌شناسی است. این ویژگی‌ها شامل، کانی‌شناسی، سنگ‌شناسی، آلتراسیون، بافت سنگ و نظایر آن است. این مرحله همواره با چالش‌های زیادی از جمله زمان‌بر و هزینه‌بر بودن، نیاز به تخصص بالا و همچنین نیروی انسانی برای تشخیص این الگوها و ویژگی‌ها همراه بوده است. یادگیری ماشین و یادگیری عمیق از جمله ابزاری هستند که در خدمت علوم زمین قرار گرفته‌اند. در این پژوهش با استفاده از معماری U-net به جدا کردن باطله و کانسنگ و شناسایی الگوی عیاری با استفاده از تصاویر گرفته شده از جعبه مغزه اقدام شد. ابتدا تصاویر استاندارد از جعبه مغزه‌ها تهیه و پیش‌پردازش‌های اولیه بر روی داده‌ها انجام شد، سپس با استفاده از بخش‌بندی باینری کانی‌های آهن جدا و برای بهینه‌سازی شبکه از روش‌های سعی و خطا استفاده گردید. در نهایت دقت مدل برای شناسایی کانسنگ آهن ۹۱ درصد بود. در ادامه برای بررسی‌های بیشتر از روش IOU استفاده شد. این روش یک معیار مناسب برای ارزیابی نهایی مدل بخش‌بندی تصاویر است که برای شناسایی کانسنگ آهن دقت مدل ۷۵ درصد به دست آمده است. در نهایت برای اعتبارسنجی نهایی مدل‌های به دست آمده، خروجی‌های به دست آمده از شبکه با مقادیر عیاری گزارش شده از آنالیز XRF یک مغزه بررسی و خطای شبکه ۹ درصد ارزیابی شد که دقت خوب مدل به دست آمده با توجه به داده‌های واقعی را نشان می‌دهد.

کلمات کلیدی

بخش‌بندی تصاویر، تصویر RGB جعبه مغزه، یادگیری عمیق، شبکه U-net.

استناد به این مقاله

فرمند، م.، فهیمی‌نیا، م.، اصغری، ا.، میرمحمدی، م. ص.؛ ۱۴۰۳؛ "تشخیص اتوماتیک کانسنگ از باطله با استفاده از تصاویر مغزه‌ها و شبکه عمیق U-Net". نشریه مهندسی منابع معدنی، دوره نهم، شماره ۴، ص ۱-۱۵.

DOI: 10.30479/jmre.2024.19041.1652



۱- مقدمه

کلاس با تصویر هایپراسپکترال، کانی‌های هر مغزه شناسایی می‌شوند [۲].

با توسعه شبکه‌های عصبی کانولوشن، پردازش‌های مربوط به داده‌های تصویری از جمله تحقیقات در مورد طبقه‌بندی و بخش‌بندی تصاویر، پیشرفت زیادی داشته است [۵]. در مقایسه با روش‌های سنتی یادگیری ماشین، CNNها پیچیدگی بیشتری دارند و می‌توانند اطلاعات تصویری سطح بالاتری را استخراج کنند [۶]. شبکه U-net یک معماری از CNN است که به طور خاص برای مسایل بخش‌بندی تصاویر استفاده می‌شود [۷]. در سال‌های اخیر از شبکه‌های بر پایه CNN و همچنین U-net در علوم مختلف از جمله زمین‌شناسی و علوم زمین استفاده‌های زیادی شده است. به عنوان مثال چن و همکاران به شناسایی و جدا کردن ذرات کانی‌های رسی از کانی‌های پس‌زمینه به وسیله شبکه U-net پرداختند و بدین منظور از تصاویر SEM استفاده کردند [۸]. زیبیدی و همکاران یک روش خودکار با استفاده از یک شبکه مبتنی بر CNN (ResNetXt-50) برای شناسایی سه نوع مختلف ترکیب سنگ را معرفی کردند. این رویکرد هنگام پیش‌بینی سنگ‌شناسی نمونه‌های مغزه جدید که قبلاً به شبکه آموزش داده نشده بود، به دقت ۹۳٫۱۲ درصد دست یافت [۹]. چن و همکاران با استفاده از تصاویر مغزه‌های حفاری و شبکه عصبی، روشی خودکار را برای شناسایی کانسنگ از باطله آهن معرفی کردند. در این مطالعه از شبکه MMDC-Unet که یک معماری بهینه شده از U-net است استفاده کردند. برای ارزیابی عملکرد مدل، نتایج به دست آمده با آنالیزهای شیمیایی مقایسه شده است که نشان‌دهنده کارایی این روش برای شناسایی کانی‌سازی آهن است [۱].

۱-۱- منطقه مورد مطالعه

در این پژوهش، داده‌های معدن آهن- مس آقاش کردخلج استفاده شد. شکل ۱ محدوده‌هایی از معدن که کانی‌سازی مس و آهن دارد را نشان می‌دهد. این محدوده با مساحتی در حدود ۶ کیلومترمربع در استان مرکزی و در غرب شهرستان ساوه واقع شده است. تیپ کانسار از نوع اسکارن و توده نفوذی مس‌دار است.

این معدن بر روی زون ساختاری ایران مرکزی و کمربند آتشفشانی ارومیه- دختر قرار گرفته است. جنس سنگ میزبان آن آندزیت، توف آندزیتی، توف ریبوداسیتی، لیتیک توف، لاتیت، داسیت و واحدهای نفوذی دیوریتی، کوارتز دیوریتی

شناسایی نحوه توزیع کانسنگ در معادن اهمیت بالایی دارد و معدن‌کاری بر اساس اطلاعات اولیه توزیع و حجم کانسنگ انجام می‌شود [۱]. یکی از مهم‌ترین روش‌ها، حفاری مغزه‌گیری است. این روش برای تعیین سیستماتیک پارامترهای زمین‌شناسی مانند سنگ‌شناسی، کانی‌شناسی و زون‌های آلتراسیون استفاده می‌شود [۲]. معمولاً لاگ‌برداری توسط زمین‌شناس انجام می‌شود که ممکن است همراه با خطای انسانی باشد و فرآیندی زمان‌بر است، بنابراین توسعه روش‌های اتوماتیک کانی‌شناسی و لاگ‌برداری از مغزه‌ها در زمینه شناسایی اطلاعات توزیع ماده معدنی در سریع‌ترین زمان و کمترین هزینه مفید است. با وجود مفید بودن روش‌های اتوماتیک لاگ‌برداری در عملیات اکتشافی این روش در حال حاضر جایگزین زمین‌شناس متخصص نیست و به عنوان یک لایه کمکی زمین‌شناسی از آن استفاده می‌شود. در سال‌های اخیر با استفاده از عکس‌های RGB و سنسورها روش‌های مختلفی مبتنی بر یادگیری ماشین و یادگیری عمیق برای شناسایی الگوهای کانی‌شناسی و زمین‌شناسی توسعه داده شده است. با توجه به این که این الگوریتم‌ها قابلیت استخراج و شناسایی ویژگی‌های سطح بالا از داده‌های تصویری را دارند، تاکنون نتایج خوبی را به همراه داشته‌اند. به عنوان مثال دستا و همکاران با استفاده از عکس‌های رنگی و با به کارگیری سنسور FTIR که نوعی حسگر فروسرخ است، روشی برای تشخیص کانی‌های دیواره تونل یک معدن زیرزمینی را توسعه دادند. در این روش از ترکیب الگوریتم‌های بدون نظارت k-means و نظارت شده حداکثر درست‌نمایی، نقشه‌بردار زاویه طیفی و روش طبقه‌بندی کمترین فاصله استفاده شده است [۳]. رامیل و همکاران برای شناسایی کانی‌های موجود در سنگ‌های آثار باستانی و مکان‌های فرهنگی از شبکه عصبی مصنوعی و عکس‌های رنگی استفاده کردند که در این مطالعه نتایج قابل قبولی برای شناسایی کانی‌های موجود در سه نوع گرانیت مختلف به دست آمد [۴]. لیو و همکاران از عکس‌های رنگی جعبه مغزه‌ها و اسکن هایپراسپکترال دستگاه Hylogger برای شناسایی کانی‌های موجود در عکس مغزه استفاده کردند. محدوده اسکن هایپراسپکترال در این دستگاه یک خط در مرکز هر مغزه است که تمام اطلاعات مغزه را نشان نمی‌دهد. در این مطالعه ابتدا با استفاده از یک روش بدون نظارت k-means تصویر هر مغزه به کلاس‌های جدا از هم تقسیم و با تطبیق هر

واحد نبوده و در حقیقت یک زون گسله است. روند کلی آن N100 است. شیب صفحه گسله گوشک نصرت متغیر و در سطح زمین نزدیک به ۶۰ تا ۸۰ درجه به سوی جنوب است. این گسل دارای مولفه امتداد لغز چپگرد و مولفه شیب لغز از نوع معکوس است که جابه‌جایی مولفه امتداد لغز آن بزرگتر و مهم‌تر به نظر می‌رسد ولی مقدار جابه‌جایی آن مشخص نیست.

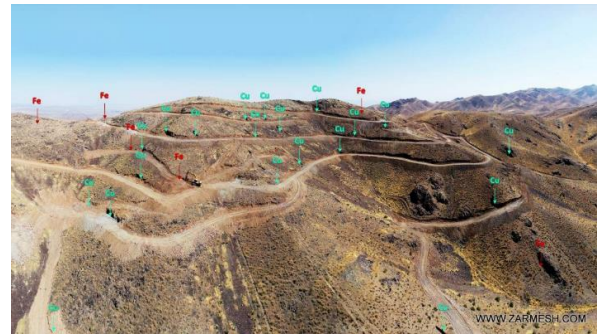
۲- روش تحقیق

روش خودکاری که برای جدا کردن کانسنگ و باطله در این تحقیق استفاده شد، بر اساس بخش‌بندی تصاویر است. مرحله اول پژوهش برداشت داده‌های تصویری از جعبه مغزه‌های حفاری و تحلیل بررسی آن‌ها با استفاده از نظر زمین‌شناس و همین‌طور نتایج آنالیز شیمیایی است. پس از بررسی‌های انجام شده بر روی داده‌ها، برچسب‌گذاری باینری برای مرحله آموزش شبکه عصبی انجام شد. برچسب‌های تولید شده مناطقی از تصاویر را نشان می‌دهد که دارای کانی‌های مگنتیت و هماتیت هستند که کانه‌زایی آهن را نشان می‌دهند. با توجه به این که هماتیت‌های موجود در مغزه‌ها هماتیت‌های خاکستری تا سیاه رنگ هستند و از لحاظ رنگی شباهت زیادی به کانی‌های مگنتیت دارند، این دو کانی هر دو به عنوان کانسنگ آهن در یک کلاس قرار داده شده‌اند.

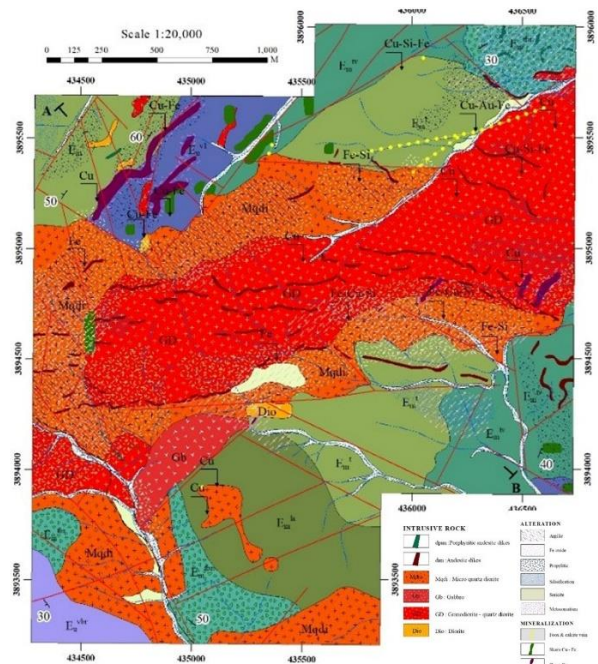
شبکه عصبی استفاده شده شبکه U-net است که یک معماری خاص از شبکه‌های عصبی کانولوشنی است. این شبکه به طور خاص برای اهداف بخش‌بندی تصاویر استفاده می‌شود. بخش‌بندی تصویر یک وظیفه در بینایی ماشین است که هدف آن تقسیم یک تصویر به بخش‌ها یا قطعات مختلف بر اساس ویژگی‌های خاصی مانند رنگ، بافت‌ها یا مرزهای اشیا است. این فرآیند برای هر پیکسل در تصویر یک برچسب یا شناسه اختصاص می‌دهد که مشخص می‌کند آن پیکسل به کدام بخش تعلق دارد [۵]. در این پژوهش از بخش‌بندی باینری استفاده شد و تصاویر به کانه و زمینه تقسیم شدند.

۲-۱- ساخت مجموعه داده‌ها

در مرحله اول برداشت داده‌ها، ابتدا جعبه مغزه‌های دارای کانسنگ آهن با توجه به نتایج آنالیز شیمیایی XRF انتخاب و سپس از این جعبه‌ها عکس‌برداری اولیه انجام شد. شکل ۳ دستگاه و شرایط عکس‌برداری برای ایزوله کردن نورهای محیطی را نشان می‌دهد. برای به دست آوردن حداکثر قدرت



شکل ۱: نمای کلی از محدوده معدنی کردخلج



شکل ۲: نقشه زمین‌شناسی ۱:۲۰۰۰۰ معدن آفاش

و گابرویی است که سن آنها بین ائوسن تا میوسن تخمین زده شده است. آلتراسیون‌های این محدوده شامل سیلیسی، سریسیتی، آرژیلیکی، پروپلیتیکی، هورنفلسی و هماتیتی است و محدوده این معدن یک زون کانه‌زایی مس-آهن به طول ۱۰۰۰ متر و عرض ۵۰۰ متر دارد. شکل ۲ نقشه زمین‌شناسی منطقه معدنی آفاش در مقیاس ۱:۲۰۰۰۰ را نشان می‌دهد که واحدهای سنگی مختلف در آن قابل مشاهده است.

محدوده اکتشافی آفاش بین دو گسل مهم گوشک نصرت و قرمز آفاش قرار گرفته است. گسل گوشک نصرت با طول گسلش بیش از ۱۵۰ کیلومتر از جنوب شرق دریاچه نمک قم آغاز و در غرب به گسل آوج می‌پیوندد. این گسل یک شکستگی

عمودی و چرخش افقی استفاده شد [۱۰] و تعداد داده‌های برچسب‌دار به ۳۰۰ افزایش یافت. از کل داده‌های ۸۰٪ برای آموزش و ۲۰٪ برای تست شبکه مورد استفاده قرار گرفت.

۲-۲- شبکه U-net

شبکه U-net یک معماری از شبکه‌های عصبی کانولوشنی است که برای وظایف بخش‌بندی تصاویر استفاده می‌شود.

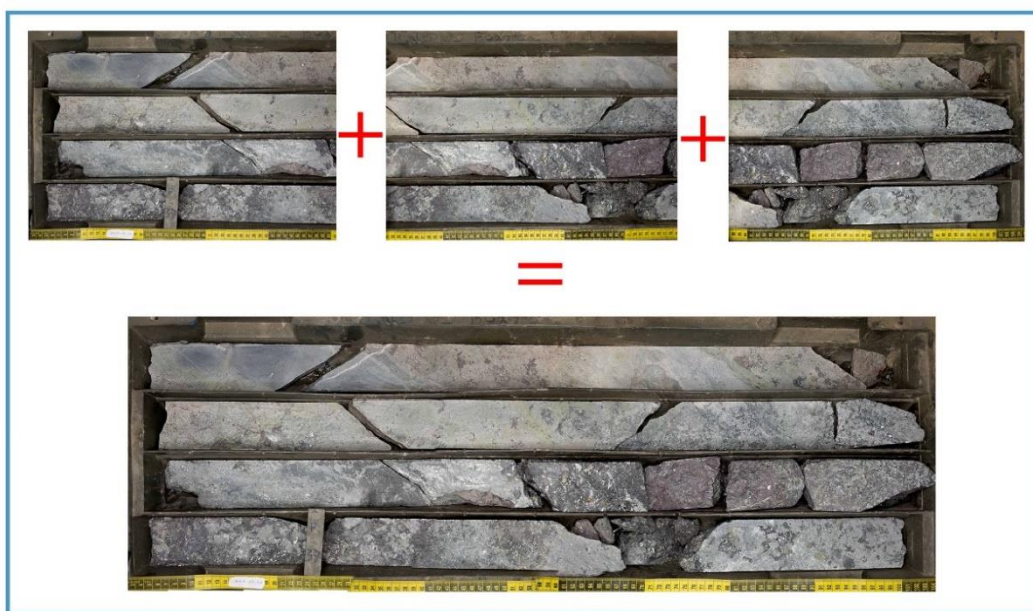


شکل ۳: نمای دستگاه استفاده شده برای عکس‌برداری از جعبه مغزه‌ها

تفکیک سه مرحله عکس‌برداری متوالی در طول جعبه انجام شد که با ادغام این سه عکس، تصویر نهایی جعبه مغزه به دست آمد که در شکل ۴ تصاویر اولیه و تصویر نهایی بعد از ادغام شدن نشان داده شده است. برای ایزوله کردن محیط عکس‌برداری و جلوگیری از ورود نورهای محیط که می‌تواند در داده‌ها نویز ایجاد کند از فون عکاسی سیاه رنگ و یک لامپ حلقه‌ای با نور سفید ثابت استفاده شد. عکس‌برداری از ۲۰ جعبه مغزه انجام شده که از ۹ گمانه مختلف بودند.

قدرت تفکیک نهایی به دست آمده عکس‌ها در حدود ۲۵۰۰×۵۸۰۰ است که در مرحله آموزش شبکه به بخش‌های کوچک‌تری تقسیم شده‌اند. برای این مرحله عکس‌های جعبه مغزه‌ها به بخش‌های کوچک‌تر ۲۵۶×۲۵۶ که برای ورودی شبکه مناسب‌تر است تقسیم شدند. عکس‌های نهایی برای آموزش از بخش‌هایی با RQD بالاتر که کانی‌ها قابل تفکیک‌تر است انتخاب شد که در نهایت ۱۰۰ عکس دارای عیار آهن به دست آمد.

برچسب‌زنی داده‌ها به روش دستی انجام و با فرمت tiff ذخیره شدند. شکل ۵ نمونه‌ای از تصاویر دارای کانسنگ آهن را همراه با برچسب آن نشان می‌دهد که در تصویر سمت چپ کانی‌های سیاه‌رنگ مگنتیت است و در تصویر سمت راست برچسب تولید شده آن با رنگ زرد مشخص شده است. برای آموزش بهتر شبکه از روش‌های داده‌افزایی کلاسیک چرخش



شکل ۴: عکس نهایی که با ادغام سه عکس‌برداری متوالی به دست می‌آید

دست و پنجه نرم می‌کنند. مانند دیگر شبکه‌های عصبی دیگر لایه ورودی حامل مقادیر پیکسل‌های عکس هستند. در طول آموزش، شبکه‌های کانولوشنی معمولا با استفاده از گونه‌ای از الگوریتم پس انتشار [۱۲] بهینه می‌شوند، جایی که گرادیان‌ها محاسبه می‌شوند و برای به‌روزرسانی وزن‌های شبکه استفاده می‌شوند. این فرآیند شامل به حداقل رساندن یک تابع هزینه است که اختلاف بین خروجی پیش‌بینی شده و مقدار حقیقی زمینه را اندازه‌گیری می‌کند.

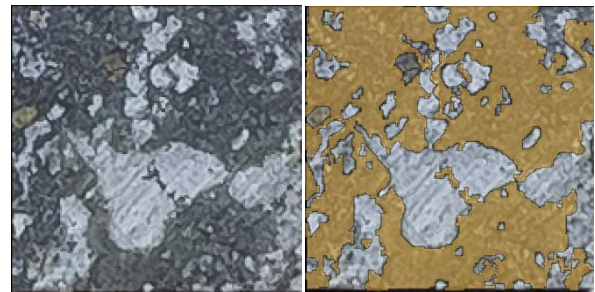
شبکه U-net یک شبکه با معماری متقارن است و از چهار بخش اصلی تشکیل شده است که شامل مسیر انقباضی (contracting path)، مسیر انبساطی (expanding path)، پل ارتباطی (bottleneck) و لایه خروجی است. شکل ۶ نمای کلی از شبکه U-net به همراه چهار بخش اصلی آن به همراه لایه‌های مختلف را نشان می‌دهد که در آن طول عمودی لایه‌ها نشان‌دهنده ابعاد تصویر و طول افقی لایه‌ها نشان‌دهنده تعداد نقشه‌های ویژگی است [۷].

۲-۲-۱- مسیر انقباضی

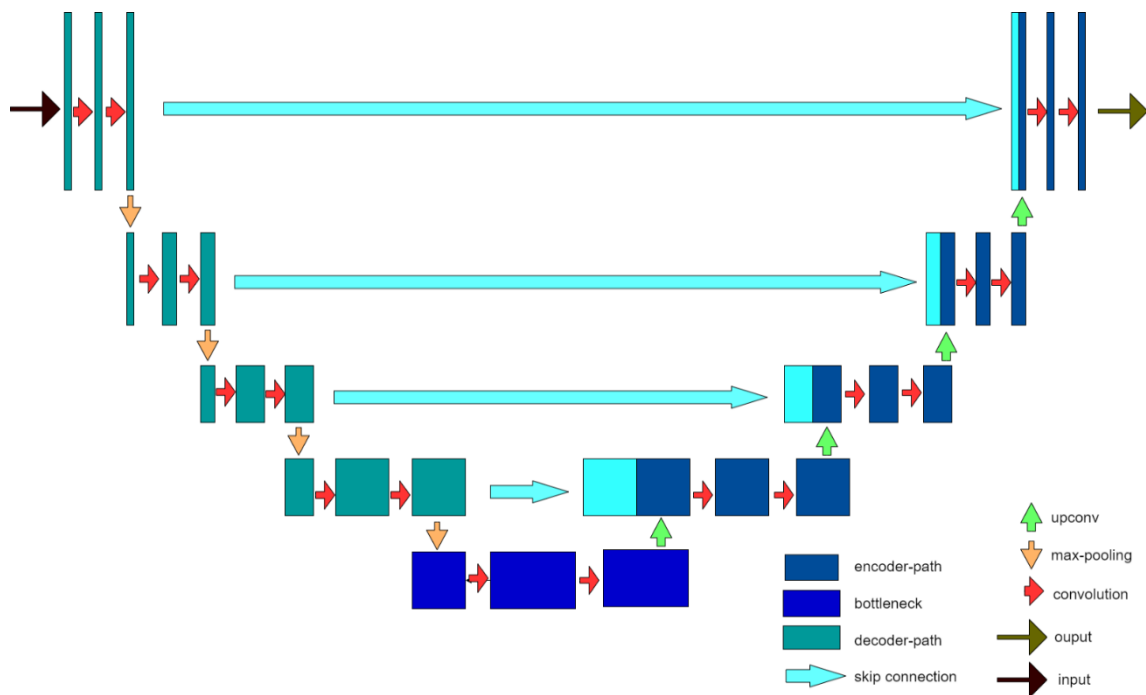
مسیر انقباضی یا رمزگذار وظیفه استخراج کانال‌های ویژگی را دارد. این مسیر از چهار بلوک رمزگذار تشکیل شده

U-net اولین بار برای بخش‌بندی عکس‌های پزشکی در دانشگاه فرایبورگ توسعه داده شد. U-net یک شبکه عصبی تماما کانولوشن است که نسبت به شبکه‌های عصبی دیگر به داده‌های کمتری نیازمند است [۷].

شبکه‌های عصبی کانولوشن (CNN) یکی از انواع مدل‌های یادگیری عمیق هستند و عمدتاً برای حل وظایف بینایی ماشین و پردازش تصویر از جمله طبقه‌بندی تصاویر، تشخیص اشیا و بخش‌بندی تصاویر استفاده می‌شود [۱۱]. یکی از بزرگ‌ترین محدودیت‌های روش‌های سنتی ANN این است که آنها با پیچیدگی محاسباتی مورد نیاز برای محاسبه داده‌های تصویری



شکل ۵: تصویر آماده شده برای آموزش آهن به همراه برچسب حقیقی



شکل ۶: معماری شبکه U-net

مکان‌یابی دقیق اشیاء و حفظ اطلاعات ریز کمک می‌کند [۷].

۲-۲-۳- گلوگاه یا پل ارتباطی

در وسط معماری U-net، لایه‌ای وجود دارد که باریک‌ترین قسمت شبکه را نشان می‌دهد. این بخش که به عنوان پل ارتباطی در بین مسیر انبساطی و انقباضی عمل می‌کند به دلیل کاهش ابعاد فضایی و تعداد کانال‌های ویژگی بیشتر، به عنوان یک ساختار گلوگاه مانند دیده می‌شود. هدف از این ساختار دریافت ضروری‌ترین اطلاعات و سطح بالاترین ویژگی‌های استخراج شده از تصویر ورودی است. این بخش با ادغام ویژگی‌ها از کانال‌هایی با قدرت تفکیک‌های مختلف می‌تواند مکان اشیاء یا قسمت‌های مختلف را تشخیص دهد و به شبکه آموزش دهد. این بخش از شبکه U-net از تعدادی لایه کانولوشن و لایه ادغامی پیشینه به همراه توابع فعال‌سازی مانند ReLU تشکیل شده است [۷].

۲-۲-۴- لایه خروجی

لایه نهایی یا خروجی شبکه وظیفه بخش‌بندی نهایی را به عهده دارد که از یک لایه کانولوشنی 1×1 تشکیل شده است و با یک تابع فعال ساز همراه است. تابع فعال ساز بسته به چند کلاسه بودن بخش‌بندی یا باینری بودن آن ممکن است متفاوت باشد. برای چندکلاسه از تابع softmax و برای باینری از sigmoid استفاده می‌شود [۱۰، ۱۱]. این لایه یک نقشه تقسیم‌بندی با ابعاد فضایی مشابه تصویر ورودی تولید می‌کند که در آن هر پیکسل نشان‌دهنده برچسب یا کلاس پیش‌بینی شده برای پیکسل مربوطه در تصویر ورودی است.

۲-۳- هاپرپارامتر

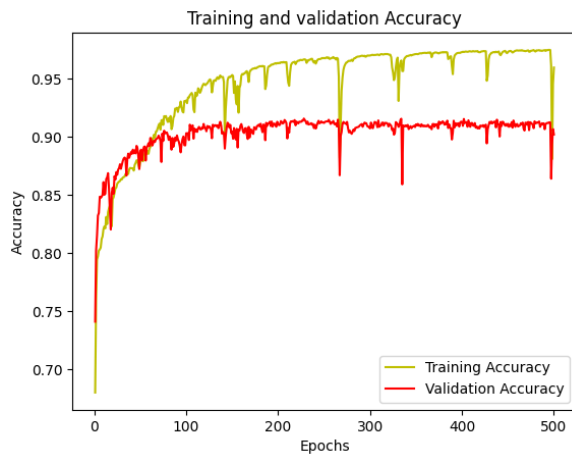
مدل‌های یادگیری عمیق CNN هاپرپارامترهای متعددی دارند که باید تعیین شوند. این پارامترها شامل پارامترهایی مانند تعداد دوره‌های آموزشی، نرخ یادگیری و بچ‌سایز که بر روی یادگیری شبکه تاثیر دارند و پارامترهایی مانند ابعاد فیلتر هستند که بر روی معماری تاثیرگذارند [۱۳]. با توجه به حساسیت شبکه‌های عصبی به مجموعه داده‌ها بهینه‌سازی هاپرپارامترهای شبکه به صورت سعی و خطا انجام شد. تعداد و ترتیب لایه‌ها و الگوریتم بهینه‌سازی بر روی پیش‌فرض شبکه U-net قرار داده شدند. به غیر از این هاپرپارامترها تعداد وسایل فیلترها، نرخ یادگیری شبکه، حذف تصادفی (Dropout) و بچ‌سایز به صورت سعی و خطا بهینه شدند.

است که هر کدام سه‌لایه دارند. دولایه کانولوشنی که بعد از هر کدام یک تابع فعال‌ساز ReLU و یک لایه pooling قرار دارد. لایه کانولوشن عملیات فیلتر کردن هوشمند را انجام می‌دهد. این لایه به وسیله محاسبه ضرب اسکالر بین مقادیر پیکسلی تصاویر ورودی و فیلتر کرنل، ورودی‌های لایه تماماً متصل برای تصمیم‌گیری نهایی را تعیین می‌کند. این عملیات شبکه را قادر می‌سازد تا الگوهای فضایی را ثبت کند و ویژگی‌هایی مانند لبه‌ها، گوشه‌ها و بافت‌ها را شناسایی کند. لایه‌های ادغامی یا pooling در هر مرحله وظیفه کاهش تعداد پارامترها را دارد و ابعاد فضایی نقشه‌های ویژگی به دست آمده از لایه‌های کانولوشن را پایین می‌آورند که این کار پیچیدگی محاسباتی را کمتر کرده و باعث می‌شود که ویژگی‌های تصاویر در ابعاد مختلف به شبکه آموزش داده شود. در هر بلوک مسیر انقباضی نسبت به بلوک قبلی تعداد کانال‌های ویژگی دو برابر می‌شوند، در حالی که ابعاد فضایی در حال نصف شدن هستند. عملیات نمونه‌برداری به طور موثر وضوح فضایی را کاهش می‌دهد و شبکه را قادر می‌سازد میدان دید بزرگ‌تری داشته باشد [۷].

۲-۲-۲- مسیر انبساطی

مسیر انبساطی یا رمزگشا برای بازیابی اطلاعات مکانی از دست رفته در مرحله رمزگذاری طراحی شده است و همین‌طور آموزش مکان‌های برچسب زده شده در هر کدام از تصاویر آموزشی به عهده این مسیر است. این مسیر از چهار بلوک رمزگشا تشکیل شده که در هر کدام یک لایه upconv وجود دارد که بعد از آن دولایه کانولوشنی با تابع فعال‌ساز ReLU وجود دارد. لایه upconv برای افزایش ابعاد تصویر و بازسازی آن استفاده می‌شود. این لایه بازسازی تصویر را با به کارگیری معکوس عملیات کانولوشن انجام می‌دهد. در هر بلوک این مسیر کانال‌های ویژگی نصف و ابعاد کانال‌ها دوبرابر می‌شوند که این به بازسازی جزئیات دقیق خروجی بخش‌بندی شده کمک می‌کند. اتصالات پرشی که یک عضو دیگر از مسیر انبساطی است یک عنصر حیاتی از U-net است. آن‌ها اتصالات بین بلوک‌های رمزگذار و رمزگشای مربوطه هستند که لایه‌های میانی را دور می‌زنند. با اتصال مستقیم لایه‌ها با ابعاد یکسان، اتصالات پرشی، رمزگشا را قادر می‌سازد تا به جزئیات دقیق‌تر از مسیر رمزگذار دسترسی پیدا کند و آن‌ها را ترکیب کند. در واقع وظیفه upsampling که به عهده لایه‌های رمزگشا بود به وسیله این اتصالات پرشی بهینه می‌شوند و در نهایت این اتصالات به

یکی از شبکه‌ها انتخاب شد. میزان دقت و خطای شبکه برای آموزش و تست در جدول ۲ ذکر شده است و در نهایت مدل ۱۱ انتخاب شد. شکل ۸ نمودار دقت و خطای نهایی را پس از حذف تاثیر وزن‌های تصادفی نشان می‌دهد. نمودارها نشان می‌دهد که شبکه به خوبی آموزش دیده و نمودارهای خطا و دقت به خوبی همگرا شده و بیش‌برازش رخ نداده است.



شکل ۷: نمودار دقت آموزش و تست شبکه

مقادیر استفاده شده برای هایپرپارامترها به صورت ۳،۴،۵،۸ برای بچ‌سایز، ۱۶،۳۲ برای تعداد فیلترهای لایه اول و ۰،۰۰۰۱ و ۰،۰۰۰۲ برای نرخ یادگیری بودند. نرخ حذف تصادفی برای هر کدام از لایه‌ها متفاوت و به ترتیب در دو بلوک اول ۰،۱، دو بلوک بعدی ۰،۲، در گلوگاه ۰،۳ و دو بلوک بعدی ۰،۲ بود. در دو بلوک آخر ۰،۱ قرار داده شد. در جدول ۱ مقادیر استفاده شده و مقدار نهایی برای هایپرپارامترها ارائه شده است.

جدول ۱: مقادیر هایپرپارامترها

مقدار نهایی	مقادیر استفاده شده	هایپرپارامتر
۳	۳،۴،۵،۸،۱۶	بچ‌سایز
۱۶	۱۶،۳۲	تعداد فیلتر لایه اول
۰،۰۰۰۲	۰،۰۰۰۰۱، ۰،۰۰۰۰۲	نرخ یادگیری
۰،۱ ۰،۲ ۰،۳ ۰،۲ ۰،۱	۰،۰۵، ۰،۱ ۰،۱۰، ۰،۲ ۰،۱۵، ۰،۳ ۰،۱، ۰،۲ ۰،۰۵، ۰،۱	حذف تصادفی در بلوک‌های رمزگذار، گلوگاه و رمزگشا به ترتیب

جدول ۲: نتایج عملکرد U-net با استفاده از وزن‌های اولیه متفاوت

مدل	خطا	خطای اعتبارسنجی	دقت اعتبارسنجی	دقت
۰	۰،۱۳	۰،۴۵	۹۶،۵۴	۰،۸۹
۱	۰،۱۹	۰،۴۷	۹۵،۹۰	۰،۸۹
۲	۰،۱۹	۰،۴۹	۹۵،۹۶	۰،۸۹
۳	۰،۲۰	۰،۴۹	۹۴،۷۹	۰،۸۹
۴	۰،۱۲	۰،۴۱	۹۶،۶۷	۰،۹۰
۵	۰،۱۲	۰،۴۲	۹۶،۵۴	۰،۸۹
۶	۰،۰۹	۰،۳۵	۹۶،۹۰	۰،۹۰
۷	۰،۱۱	۰،۴۰	۹۶،۷۱	۰،۹۰
۸	۰،۱۶	۰،۴۷	۹۵،۳۶	۰،۹۰
۹	۰،۱۲	۰،۴۲	۹۶،۵۱	۰،۹۰
۱۰	۰،۱۱	۰،۴۰	۹۶،۶۵	۰،۸۹
۱۱	۰،۰۸	۰،۳۸	۹۶،۸۴	۰،۹۱
۱۲	۰،۱۴	۰،۴۳	۹۵،۵۳	۰،۹۰
۱۳	۰،۱۰	۰،۴۱	۹۶،۹۷	۰،۸۹
۱۴	۰،۱۹	۰،۴۵	۹۴،۹۹	۰،۸۹

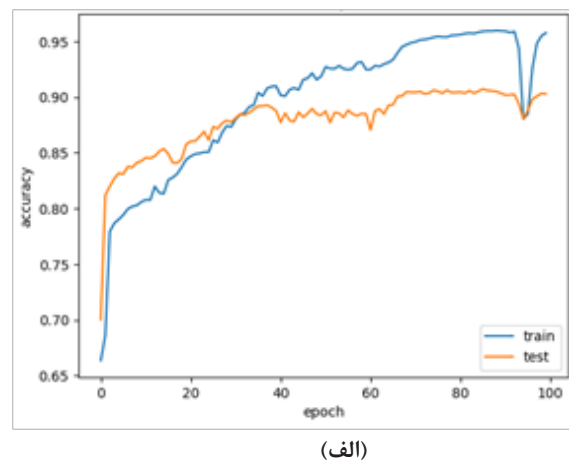
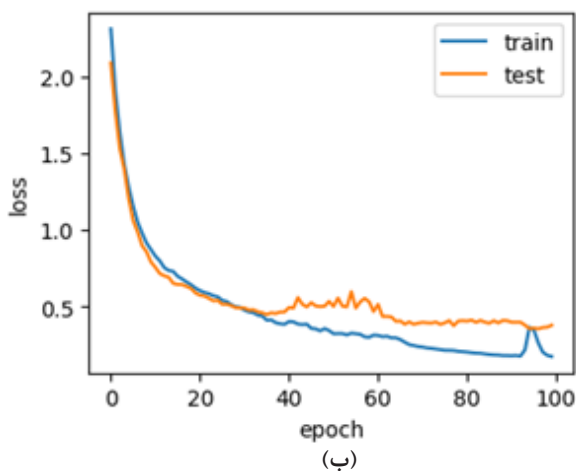
همین‌طور برای جلوگیری از بیش‌برازش از نرمال‌ساز L2 استفاده شد. این روش در مسایل یادگیری ماشین و یادگیری عمیق برای جلوگیری از بیش‌برازش و بهبود تعمیم استفاده می‌شود. این امر با اضافه کردن یک جریمه به تابع هزینه حین آموزش به دست می‌آید و به کوچک نگه‌داشتن وزن‌ها به وسیله مدل و ایجاد تعادل بین برازش‌های داده‌های آموزشی کمک می‌کند [۱۴]. شکل ۷ اجرای اولیه شبکه را نشان می‌دهد که برای پیدا کردن تعداد دوره آموزشی بهینه استفاده شد که تعداد دوره‌های آموزشی (epoch) برای شروع روی ۵۰۰ قرار گرفت تا نقطه همگرایی شبکه پیدا شود.

در نهایت با توجه به نمودارهای دقت و خطای به دست آمده که در شکل ۷ نمودار زرد دقت تست و نمودار قرمز دقت آموزش شبکه را نشان می‌دهد و بررسی آنها مشخص شد که شبکه بعد از ۱۰۰ دوره آموزشی همگرا می‌شود و همه پارامترها به خوبی آموزش دیده‌اند و شبکه به آموزش بیشتر نیاز ندارد. پس از یافتن تعداد دوره‌های بهینه برای جلوگیری از تاثیر وزن‌های اولیه تصادفی و ایجاد خطا، شبکه ۱۵ بار اجرا شد. بر اساس مقدار دقت پیش‌بینی، خطای پایین تست و آموزش و پارامتر IOU که در بخش ۳ معرفی شده است،

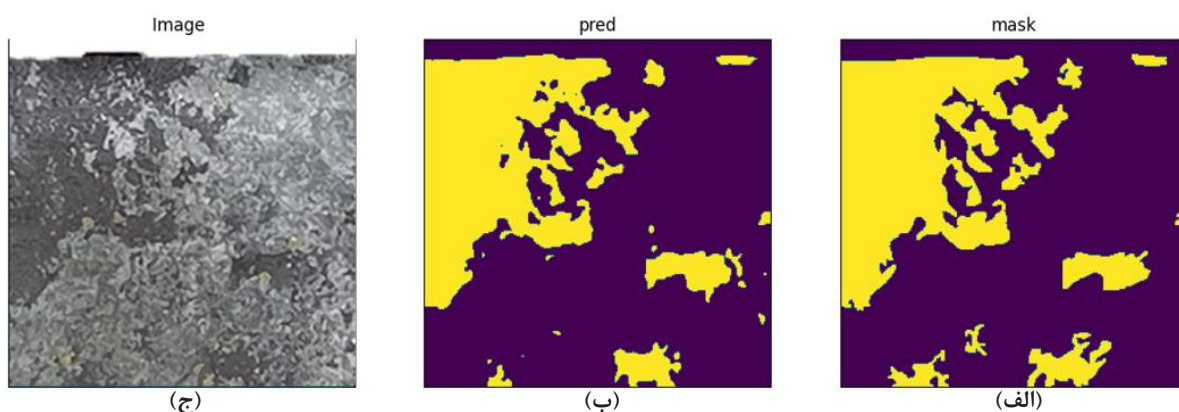
۳- نتایج و ارزیابی

بردن قدرت تفکیک تصاویر برداشت شده و استفاده از دوربین بهتر و یا پیچیده تر شدن شبکه به بهبود نتایج کمک می کند. سنجش کمی عملکرد شبکه به وسیله پارامترهای خروجی سنجیده می شود که در این جا دقت یا همان Accuracy به عنوان یکی از معیار سنجش قرار داده می شود و دقت شبکه در حدود ۹۱٪ بود، اما در مسایل بخش بندی معیار اصلی سنجش پارامتری به نام IOU (Intersection Over Union) است [۱۵]. علت آن این است که برای محاسبه دقت در یک مساله بخش بندی تصاویر منطبق بودن پیش بینی شبکه و برچسب حقیقی برای همه کلاسها تاثیرگذار است و تشخیص خوب پس زمینه تصویر دقت شبکه را بالا می برد، اما در IOU فقط انطباق پیش بینی و برچسب برای یک کلاس اهمیت دارد، بنابراین در سنجش عملکرد شبکه برای تشخیص کانسنگ آهن، این پارامتر معیار بهتری است. IOU یک معیار ارزیابی رایج در بینایی کامپیوتر است، به ویژه در بخش بندی تصاویر که برای اندازه گیری دقت پیش بینی های یک الگوریتم با مقایسه همپوشانی بین مناطق مرزی پیش بینی شده و برچسب های حقیقی استفاده می شود [۱۶]. این پارامتر به دست آوردن نسبت مساحت همپوشانی بین مناطق پیش بینی شده و برچسب های حقیقی به اجتماع این دو ناحیه به دست می آید. در مرحله نهایی IOU به دست آمده برای کانی های آهنی در حدود ۷۵٪ به دست آمد که نشان دهنده شناسایی بیشتر کانی های آهن است.

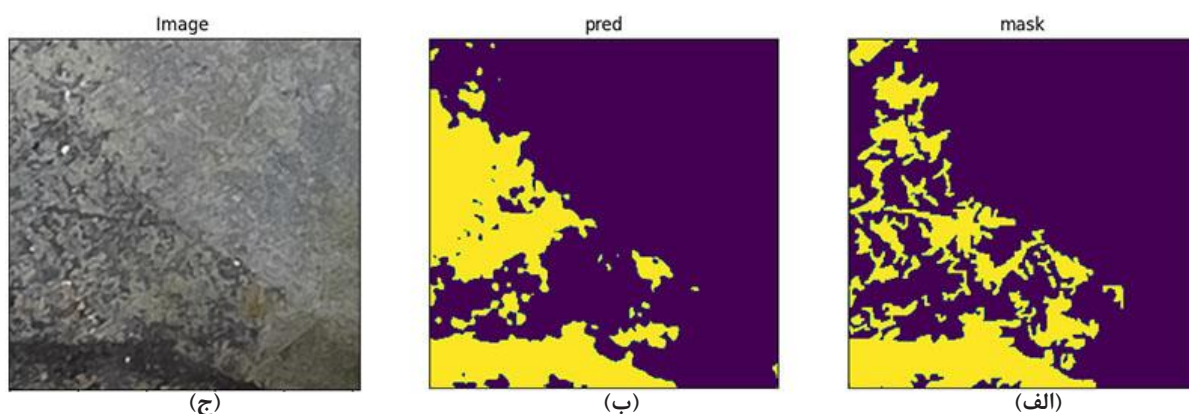
کدهای شبکه عصبی U-net و مراحل بخش بندی تصاویر در زبان برنامه نویسی پایتون انجام شد و از نرم افزار جانبی دیگری استفاده نشد. همچنین سخت افزار استفاده شده برای اجرای برنامه پایتون شامل پردازشگر Ryzen7 GTX 1650، رم 16GB و پردازنده گرافیکی 3750H استفاده شد. همچنین نسخه محیط GPU tensorflow و نسخه Python3.9 برای ساخت شبکه عصبی U-net استفاده شدند. خروجی شبکه تصاویر بخش بندی شده ای هستند که بخش های مشخص شده در آن نشان دهنده کانی سازی آهن هستند. برای سنجش نحوه عملکرد شبکه به صورت کیفی تصاویر برچسب زده شده، تصاویر اصلی و تصاویر خروجی شبکه با هم مقایسه شدند. شکل های ۹ و ۱۰ نمونه هایی از عملکرد دقیق و کم دقت شبکه را نشان می دهد. شکل ۹ نمونه ای از پیش بینی های درست شبکه را نشان می دهد که در آن مناطق دارای آهن پیش بینی شده با برچسب حقیقی مطابقت دارد. شکل ۱۰ نمونه ای از پیش بینی های نه چندان درست شبکه است که به دلیل دانه بندی ریز کانی ها شبکه پیش بینی کاملاً منطقی با برچسب حقیقی نداشته است. با مشاهده عکس ها و مقایسه آن ها می توان دریافت که شبکه در بخش هایی به غیر از بخش هایی که کانی ها بسیار ریزدانه اند و یا مغزه دارای شکستگی است، پیش بینی درستی از محل وجود کانسنگ آهن می کند و به خوبی آنها را تشخیص می دهد. بالا



شکل ۸: الف) نمودار دقت داده های آموزشی و تست پس از حذف تاثیر وزن های تصادفی و ب) نمودار خطای داده های آموزشی و تست پس از حذف تاثیر وزن های تصادفی



شکل ۹: نمونه‌ای از پیش‌بینی دقیق شبکه الف) تصویر برچسب‌زده شده، ب) تصویر خروجی نهایی و ج) تصویر اولیه



شکل ۱۰: نمونه‌ای از پیش‌بینی نه چندان دقیق شبکه الف) تصویر برچسب‌زده شده، ب) تصویر خروجی نهایی و ج) تصویر اولیه

درصد پیکسل‌هایی که به عنوان آهن تشخیص داده شده‌اند و درصد کانی‌های آهن در یک طول مشخص است. مرحله پایانی این پژوهش اعتبارسنجی عیارهای شبکه به وسیله داده‌های عیاری آنالیز XRF است. با انجام این اعتبارسنجی مفید بودن این روش اتوماتیک بر پایه شبکه عصبی و بخش‌بندی تصاویر، سنجیده می‌شود. با وجود این که مگنتیت (یا کانی‌های دیگری که با این روش شناسایی می‌شوند) به صورت ۱۰۰ درصد و خالص نیستند، اما با توجه به این که کانی‌های با خلوص بسیار پایین در این روش شناسایی نمی‌شوند، درصد پیش‌بینی شده به مقدار واقعی نزدیک و قابل استناد است. علاوه بر این، داده‌های به دست آمده از این روش برای استفاده در مراحل اولیه اکتشافی است و جایگزین روش‌های آنالیز شیمیایی نیست.

روش کلی کار به اینصورت است که ابتدا مغزه‌هایی که عیار آن‌ها را در یک متر گزارش کرده‌اند انتخاب شود، سپس

$$IOU = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

که در آن:

A: مناطق مرزی پیش‌بینی شده

B: مناطق مرزی برچسب‌زده شده است.

۳-۱- اعتبارسنجی شبکه به وسیله داده‌های عیاری

اعتبارسنجی پیش‌بینی‌های یک پروژه بخش‌بندی تصاویر به روش‌های مختلفی انجام می‌شود. از جمله آنها مقایسه کردن پیش‌بینی‌های داده‌های تست یا همان داده‌های اعتبارسنجی با برچسب‌های مرجع حقیقی است که این کار در مراحل قبلی انجام شد. این مقایسه با به دست آوردن IOU یا مقایسه چشمی نتایج انجام و به صورت کیفی و کمی نتیجه‌گیری می‌شود.

یکی از روش‌های اعتبارسنجی نتایج پیش‌بینی‌های شبکه که در این پژوهش امکان اجرای آن فراهم است، مقایسه بین

موفقیت آمیز بودن معماری U-net در جداسازی کانسنگ از باطله و همین طور توانایی ثبت جزئیات کوچک به وسیله آن را نشان داد. با این وجود عملکرد شبکه در بخش‌های با شکستگی و درزه‌های کمتر، بهتر است. نتایج شبکه نشان می‌دهد که مدل آموزش دیده بر روی کانی‌های آهنی می‌تواند بخش‌هایی از مغزه‌ها که دارای کانی‌های مگنتیتی و هماتیتی هستند، RQD بالایی دارند و دارای شکستگی کمتری هستند را با دقت مناسبی پیش‌بینی کند. از این مدل برای تشخیص کانسنگ و باطله در تصاویر جعبه‌مغزه و مغزه‌های کانسارهای مشابه نیز می‌توان استفاده کرد. به وسیله این روش با تشخیص کانی‌های آهن در طول‌های مشخص می‌توان برآورد اولیه‌ای از عیار در کانسنگ داشت و از آن به عنوان یک لایه اکتشافی که سریع و با هزینه کم تولید می‌شود، استفاده کرد.

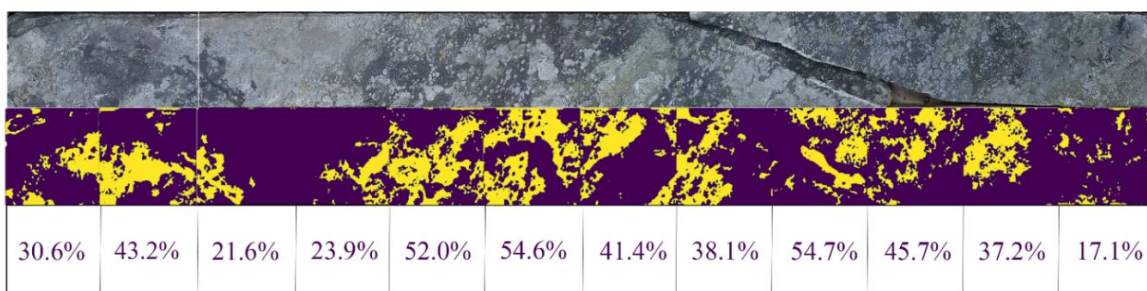
استفاده از بخش‌بندی تصویر ابزار مفیدی برای زمین‌شناسان و مهندسان معدن فراهم می‌کند که مغزه‌های حفاری را به صورت بصری، بررسی و تجزیه و تحلیل کنند. علاوه بر آن روش به کار گرفته شده در این پژوهش نسبت به روش دستی، سرعت و دقت بالاتری دارد و تاثیر خطای انسانی و تفسیر ذهنی را کاهش می‌دهد. با توجه به مقایسه انجام شده بین پیش‌بینی‌ها و برچسب‌های حقیقی مشخص شد که IOU به عنوان یک معیار سنجش، برای تشخیص خوب یا بد بودن عملکرد شبکه در جداسازی کانسنگ و باطله معیار مناسب‌تری نسبت به دقت است. با توجه به این که در این پژوهش از برچسب‌گذاری دستی و با نظر زمین‌شناس استفاده شد، برای بهبود نتایج و سرعت بخشی به اجرای روش، از برچسب‌گذاری با استفاده از روش‌های بدون نظارت مانند k-means استفاده کرد. برچسب‌های اتوماتیک بسته به داده‌ها، نتایج بهتری نسبت به برچسب دستی دارند. همچنین برای کم کردن تاثیر توزیع نامتعادل بین کلاس‌ها از معیار سنجش مانند f1score

مانند مرحله آماده‌سازی داده‌ها، تصویر آن بخش مغزه به ابعاد مورد نظر 256×256 برای ورودی شبکه تبدیل شود. بدون نیاز به برچسب‌زنی تصاویر در این مرحله تصاویر به شبکه آموزش دیده‌ای که پارامترها و هایپر پارامترهای آن بهینه شده‌اند، داده شد. خروجی شبکه طبق روند قبلی عکس‌های بخش‌بندی شده‌ای هستند که کانی‌های آهنی تصاویر در آن جدا شده است.

با کنار یکدیگر قراردادن عکس‌ها و به دست آوردن میانگین پیکسل‌هایی که در آن کانی‌های آهنی وجود دارند و مقایسه با داده عیاری، می‌توان معتبر بودن یا نبودن عملکرد شبکه را تایید کرد. شکل ۱۱ نشان‌دهنده یکی از مغزه‌هایی است که به وسیله این روش اعتبارسنجی شدند. با نگاه اولیه به این تصویر و مقایسه پیش‌بینی‌های شبکه با مناطق دارای کانی‌سازی آهن، شناسایی بخش‌های دارای کانی مگنتیت تایید می‌شود. علاوه بر آن با توجه به نسبت پیکسل‌های مگنتیت به کل تصویر اعتبارسنجی کمی نیز انجام شده است. میانگین درصد‌های گزارش شده از کانی مگنتیت در هر کدام از تصاویر کوچک 256×256 ، 38.52% است که با توجه به درصد جرمی آهن نسبت به مگنتیت (72.36%)، درصد آهن توتال، 27.87% به دست می‌آید که با مقایسه با درصد آنالیز XRF که 30.83% است، می‌توان به دقت قابل قبول مدل پی برد. با توجه به این که مدل شبکه عصبی تصاویر را در دو بعد آموزش می‌بیند و عیار گزارش شده با توجه به آنالیز شیمیایی یک مغزه نیم استوانه‌ای است، خطای ۹ درصدی به دست آمده قابل قبول است.

۴- نتیجه‌گیری

در این پژوهش روشی خودکار برای جداسازی کانسنگ و باطله در تصاویر جعبه‌مغزه‌های حفاری، بر اساس بخش‌بندی تصاویر و شبکه U-net، ارائه شد. خروجی‌های نهایی شبکه



شکل ۱۱: مقایسه پیش‌بینی شبکه و تصویر اصلی و اعتبارسنجی نتایج

- MICCAI 2015, Lecture Notes in Computer Science, Vol. 9351, Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-24574-4_28.
- [8] Chen, Z., Liu, X., Yang, J., Little, E., and Zhou, Y. (2020). "Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin". *Computers & Geosciences*, 138: 104450. DOI: 10.1016/j.cageo.2020.104450.
- [9] Alzubaidi, F., Mostaghimi, P., Swietojanski, P., Clark, S. R., and Armstrong, R. T. (2021). "Automated lithology classification from drill core images using convolutional neural networks". *Journal of Petroleum Science and Engineering*, 197: 107933. DOI: 10.1016/j.petrol.2020.107933.
- [10] Shorten, C., and Khoshgoftaar, T. M. (2019). "A survey on Image Data Augmentation for Deep Learning". *Journal of Big Data*, 6: 60. DOI: 10.1186/s40537-019-0197-0.
- [11] Lecun, Y., Bengio, Y., and Hinton, G. (2015). "Deep learning". *Nature*, 521(7553): 436-444., DOI: 10.1038/nature14539.
- [12] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). "Learning representations by back-propagating errors". *Nature*, 323(6088): 533-536. DOI: 10.1038/323533a0.
- [13] Lee, Y., Sim, W., Park, J., and Lee, J. (2022). "Evaluation of Hyperparameter Combinations of the U-Net Model for Land Cover Classification". *Forests*, 13(11): 1813. DOI: 10.3390/f13111813.
- [14] van Laarhoven, T. (2017). "L2 Regularization versus Batch and Weight Normalization". 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, arXiv, 1-9. DOI: <https://doi.org/10.48550/arXiv.1706.05350>.
- [15] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., and Terzopoulos, D. (2022). "Image Segmentation Using Deep Learning: A Survey". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7): 3523-3542. DOI: 10.1109/TPAMI.2021.3059968.
- [16] Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., and Zisserman, A. (2010). "The pascal visual object classes (VOC) challenge". *International Journal of Computer Vision*, 88(2): 303-338. DOI: 10.1007/s11263-009-0275-4.
- [17] Kaiming, H., Gkioxari, G., Dollár, P., and Girshick, R. (2017). "Mask R-CNN". *IEEE Access*, 2961-2969.
- به همراه IoU می‌توان استفاده کرد. همین‌طور معماری‌های دیگر مناسب برای بخش‌بندی تصاویر مانند mask R-CNN ممکن است مفید باشد [۱۷]. این شبکه یکی از نسخه‌های faster R-CNN (region based CNN) است که برای وظایف تشخیص اشیا و بخش‌بندی تصاویر استفاده می‌شود. برای این کار شبکه از یک سری مناطق پیشنهادی بالقوه استفاده می‌کند و با محدود کردن آنها تشخیص نهایی را انجام می‌دهد.

۵- مراجع

- [1] Jin, C., Wang, K., Han, T., Lu, Y., Liu, A., and Liu, D. (2022). "Segmentation of ore and waste rocks in borehole images using the multi-module densely connected U-net". *Computers & Geosciences*, 159: 105018. DOI: 10.1016/j.cageo.2021.105018.
- [2] Liu, J., Chen, W., Muller, M., Chalup, S., and Wheeler, C. (2019). "An automatic HyLoggerTM mineral mapping method using a machine-learning-based computer vision technique". *Australian Journal of Earth Sciences*, 66(7): 1063-1073. DOI: 10.1080/08120099.2019.1600167.
- [3] Desta, F., and Buxton, M. (2017). "The use of RGB Imaging and FTIR Sensors for Mineral mapping in the Reiche Zeche underground test mine, Freiberg". *Real Time Mining, Conference on Innovation on Raw Material Extraction, Amsterdam*, 103-127.
- [4] Ramil, A., López, A. J., Pozo-Antonio, J. S., and Rivas, T. (2018). "A computer vision system for identification of granite-forming minerals based on RGB data and artificial neural networks". *Measurement*, 117: 90-95. DOI: 10.1016/j.measurement.2017.12.006.
- [5] Cheng, H. D., Jiang, X. H., Sun, Y., and Wang, J. (2001). "Color image segmentation: Advances and prospects". *Pattern Recognition*, 34(12): 2259-2281. DOI: 10.1016/S0031-3203(00)00149-7.
- [6] Varatharasan, V., Shin, H. S., Tsourdos, A., and Colosimo, N. (2019). "Improving Learning Effectiveness for Object Detection and Classification in Cluttered Backgrounds". 2019 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED UAS), 25-27 Nov., 78-85. DOI: 10.1109/REDUAS47371.2019.8999695.
- [7] Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: Navab, N., Hornegger, J., Wells, W., and Frangi, A. (Eds.), *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015*.