



Smart Variogram Modeling using Deep Learning Method

Shokufeh Monavvari¹, Mohammad Fahiminia², Omid Asghari^{3*}

1. Department of Mining Engineering, Collage of Engineering University of Tehran, Iran,
shokufeh.monavari@ut.ac.ir

2. Department of Mining Engineering, Collage of Engineering University of Tehran, Iran,
m.fahiminia@ut.ac.ir

3. Department of Mining Engineering, Collage of Engineering University of Tehran, Iran,
o.asghari@ut.ac.ir

Received: 2022/10/24 - Accepted: 2023/06/13

Abstract

Calculating variograms and spatial continuity is one of the first and most essential processes in geostatistical modeling, which is a long and experience-oriented process. This article presents an intelligent variogram modeling method using deep learning that can increase the speed of variogram modeling and prevent common errors in manual variogram model fitting. In this method, two convolutional neural networks are used. The first CNN converts the initial data into a 2D simulated map based on various variogram models. The output of this model is entered into the second convolutional neural network as input, and the variogram parameters are predicted. The accuracy of the proposed model for synthetic 2D data was 97 %. Also, the accuracy of the obtained model for Nouchon area geochemical data compared to manual fitting was 90%.

Keywords

Variogram model, deep learning, convolutional neural network, Geostatistics.

* Corresponding Author



1- Introduction

In geostatistical modeling, spatial continuity is a fundamental and vital step intricately linked to the variogram. Calculating experimental variograms and traditional variogram modeling is time-consuming and requires a high level of experience. Consequently, interpreting and fitting the appropriate model are always the main challenges in this field. Acknowledging these challenges, the automation of variogram modeling has been extensively debated, leading to the introduction of various models [1]. One of the standard methods of calculating the variogram model is the weighted least squares method. This method minimizes the distance of each point of the experimental variogram with the fitted model [2]. Although the LS method has a simple function, in some cases, the optimal model fitted by this method cannot be the best variogram model [3]. The following method in automatic variogram fitting is the maximization method (ML) [4]. In this approach, the parameters of the variogram model are estimated by minimizing the negative logarithm of the likelihood function. Still, this method is effective only when data follows a multivariate Gaussian distribution. Also, in all of the presented, the fitted variogram model is based on an experimental variogram. This paper presents an automated variogram modeling framework, which estimates variogram parameters independently of the experimental variogram using a deep learning method. The proposed approach increases speed, prevents errors in manual fitting, and reduces the subjectivity of variogram modeling.

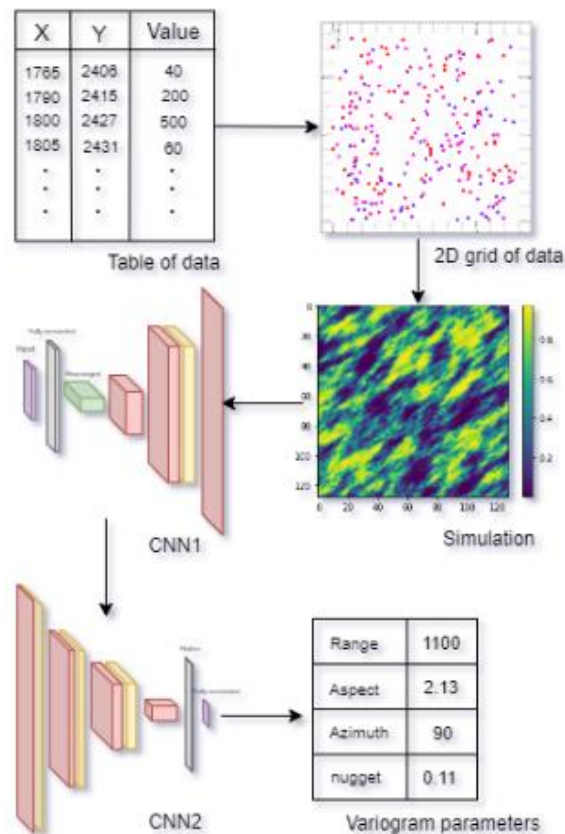


Figure 1- General flowchart of the algorithm

2- Methods

This method consists of two convolutional neural networks. The model's procedure for predicting variogram parameters is presented in Figure 1. The training dataset for the model comprises three data series, including 128 points randomly distributed within a 128x128 grid, simulated images with the SGS algorithm, and corresponding variogram parameters. New data is extracted from the simulations and coordinated with the spatial positions of the initial data, which is then utilized as input for the first CNN network. Before entering the network, this dataset is split

into three parts: 65% for training, 15% for validation, and the remaining 20% for model testing. The first CNN network converts the input data into a 2D simulated map through one MLP layer, five convolution layers, and four upsampling layers. For this purpose, training the first network with the input data and their corresponding simulations is necessary. The output of this model is entered into the second convolutional neural network as input, and the variogram parameters (including range, azimuth, ratio, and nugget effect) are predicted, using one MLP layer, 8 convolution layers, 4 Max pooling layers, and one Dropout layer. The training process for these CNNs includes minimizing the Mean Squared Error (MSE) using the Adam optimization method, with a learning rate 0.001 for both models.

3- Findings and Argument

First, the proposed algorithm is implemented on synthetic 2D data, and the parameters of the CNN models are optimized. The network achieved optimal results by integrating ReLU activation functions in convolutional layers and Sigmoid activation functions in fully connected layers, with an optimized BatchSize set to 64. The accuracy of the proposed model was 97%. At the same time, each predicted parameter was assessed individually. The Nugget Effect showed the highest accuracy at 98%, and the Ratio, Range, and azimuth were predicted with an accuracy of 96%, 95%, and 88%, respectively. Then, the proposed algorithm is used for variogram modeling of Nouchon area geochemical data, including Cu, Zn, and Pb. The Nugget Effect, Ratio, Azimuth, and Range accuracies stand at 99%, 90%, 92%, and 95%, respectively. Also, the obtained model is compared with the manually fitted model by an expert, and the accuracy of its parameters for the Azimuth, Range, Ratio, and Nugget Effect Range is determined to be 96%, 90%, 84%, and 78% based on expert models criteria.

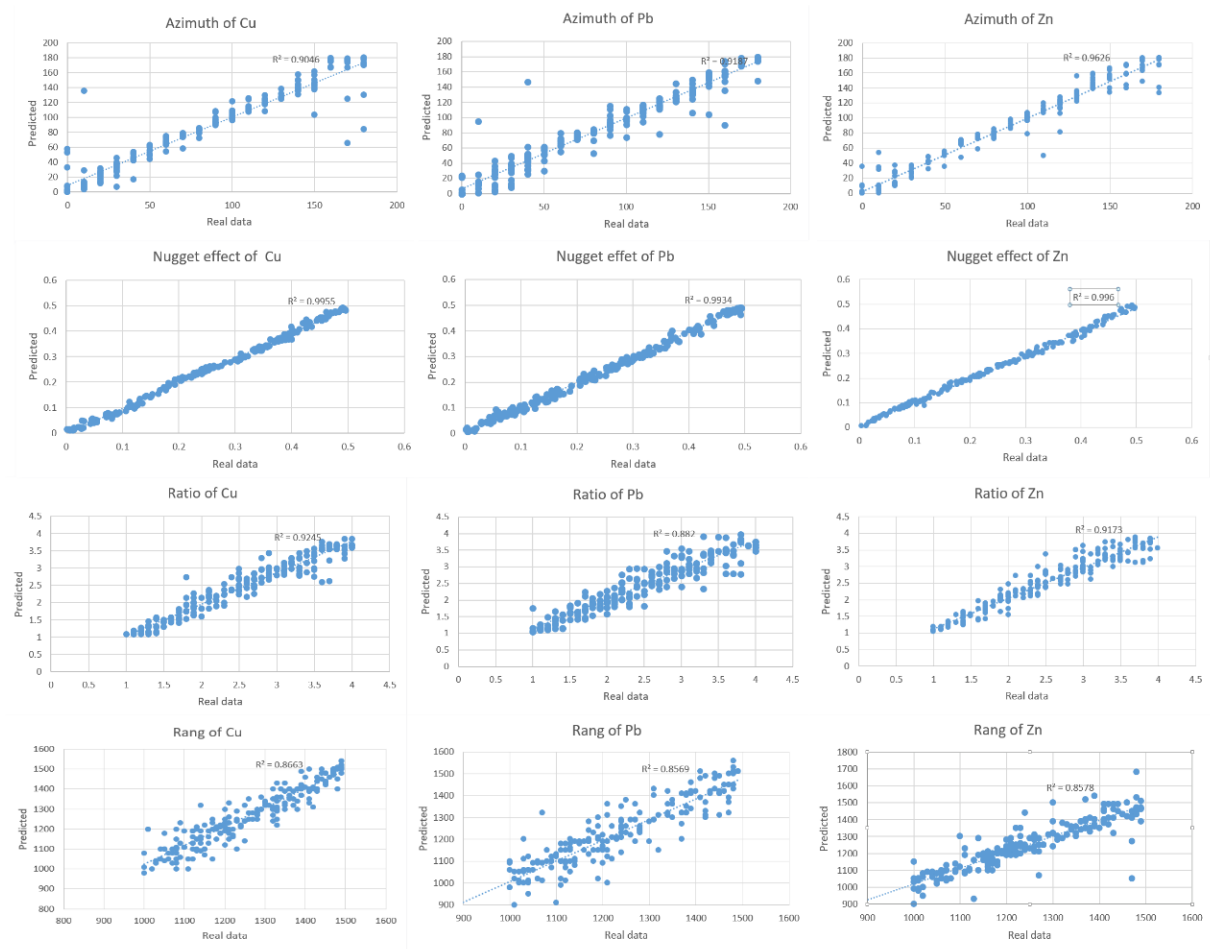


Figure 2- Validation of predicted variogram model parameters for Data of the Nochun region

4- Conclusions

This paper introduces an intelligent modeling approach utilizing deep learning, comprising two CNN networks. CNN1 transforms initial data into 2D images corresponding to the simulations conducted. Subsequently, the output of this network enters the CNN2 network, which extracts spatial relationships and translates them into variogram parameters. As illustrated in Figure 2, the proposed network demonstrates high accuracy in determining variogram parameters, including 99% for the nugget effect, 90% for the ratio, 86% for the range, and 92% for azimuth. A significant drawback of this approach is its limitation to two-dimensional data. Also, in the proposed algorithm, the fitted variogram model is a single structure but can be adapted for multi-structure models by increasing the parameters of the model.

References

1. H. Jo and M. J. Pyrcz, "Automatic Semivariogram Modeling by Convolutional Neural Network," *Math. Geosci.*, vol. 54, no. 1, pp. 177–205, 2022, doi: 10.1007/s11004-021-09962-w.
2. C. A. Gotway, "Fitting semivariogram models by weighted least squares," *Comput. Geosci.*, vol. 17, no. 1, pp. 171–172, 1991, doi: 10.1016/0098-3004(91)90085-R.
3. M. E. Rossi and C. V. Deutsch, *Mineral resource estimation*. 2014. doi: 10.1007/978-1-4020-5717-5.
4. K. V. Mardia and R. J. Marshall, "Maximum likelihood estimation of models for residual covariance in spatial regression," *Biometrika*, vol. 71, no. 1, pp. 135–146, 1984, doi: 10.1093/biomet/71.1.135.

مقاله پژوهشی

مدلسازی هوشمند واریوگرام با استفاده از یادگیری عمیق

شکوفه منوری^۱، محمد فهیمی نیا^۲، امید اصغری^{۳*}

۱. دانشجوی کارشناسی مهندسی معدن؛ دانشگاه تهران، shokufeh.monavari@ut.ac.ir

۲. دانشجوی دکتری مهندسی معدن؛ دانشگاه تهران، m.fahiminia@ut.ac.ir

۳. دانشیار دانشکده مهندسی معدن، دانشگاه تهران، o.asghari@ut.ac.ir

دریافت: ۱۴۰۱/۰۸/۰۲ - پذیرش: ۱۴۰۲/۰۳/۲۳

چکیده

محاسبه واریوگرام و پیوستگی فضایی یکی از اولین و مهم‌ترین فرآیندها در مدل‌سازی زمین‌آماري بوده که فرآیندی زمان‌بر و تجربه‌محور است. همچنین به دلیل پیچیدگی‌های محاسبه واریوگرام تجربی، تفسیر و برازش مدل مناسب همواره یکی از چالش‌های اصلی در این زمینه است. در این مقاله یک روش مدل‌سازی هوشمند واریوگرام با استفاده از یادگیری عمیق ارائه شده است که می‌تواند سرعت برازش مدل واریوگرام را افزایش دهد و مانع بروز خطاهای متداول در برازش دستی مدل واریوگرام شود. در این روش از دو شبکه عصبی کانولوشن استفاده شده است. شبکه اول داده‌های اولیه را تبدیل به نقشه دوبعدی شبیه‌سازی شده بر مبنای مدل‌های مختلف واریوگرام می‌کند. بدین منظور نیاز است تا شبکه اول، با داده‌های اولیه و شبیه‌سازی‌های مانند آن‌ها آموزش داده شود؛ سپس خروجی این مدل وارد شبکه عصبی کانولوشن دوم شده که در این شبکه تصاویر دوبعدی شبیه‌سازی شده به‌عنوان ورودی به شبکه داده می‌شود و پارامترهای واریوگرام شامل دامنه، آزیموت جهت اصلی، نسبت دامنه جهت اصلی به جهت فرعی و اثر قطعه‌ای پیش‌بینی می‌شود. در این مقاله ابتدا الگوریتم پیشنهادی بر روی داده‌های دوبعدی مصنوعی پیاده‌سازی و پارامترهای مدل بهینه شده است. دقت مدل در پیش‌بینی پارامترهای واریوگرام ۹۷ درصد بوده است. سپس از الگوریتم پیشنهادی برای مدل‌سازی واریوگرام داده‌های ژئوشیمیایی منطقه نوجون که شامل عناصر Cu، Zn و Pb استفاده شد که دقت مدل واریوگرام به‌دست‌آمده نسبت به مدل دستی برازش شده ۹۰ درصد است.

کلمات کلیدی

مدل واریوگرام، یادگیری عمیق، شبکه عصبی کانولوشن، زمین‌آمار

۱- مقدمه

می‌شود زیرا برای برازش مدل واریوگرام نیاز به محاسبه واریوگرام‌های تجربی است [۹].

اولیور و وبستر^۲ استدلال کردند در صورتی که رویکرد LS آگاهانه استفاده شود، می‌تواند نتایج رضایت‌بخشی را در ۹۰ درصد مواقع به همراه داشته باشد [۱۰] اما در برخی موارد بهینه‌ترین مدل برازش‌شده با روش LS نمی‌تواند بهترین مدل واریوگرام باشد. به عنوان مثال برای بدست آوردن اثر قطعه‌ای، از واریوگرام قائم (واریوگرام در راستای گمانه‌ها) استفاده می‌شود و سپس با استفاده از اثر قطعه‌ای بدست آمده، مدل واریوگرام در جهات دیگر برازش می‌شود. این روش با وجود اینکه مقدار حداقل مربعات را کمینه نمی‌کند، اغلب مدل واریوگرام صحیح‌تری را ارائه می‌کند [۵].

روش بعدی در برازش اتوماتیک مدل واریوگرام روش بیشینه درست‌نمایی (ML) است. این روش پارامترهای مدل واریوگرام را با کمینه کردن تابع منفی لگاریتم داده‌های خام تحت فرض چند گاوسی تخمین می‌زند [۱۲-۱۰]. از معایب این روش می‌توان به وابسته بودن پارامترهای مدل برازش شده به فرضیه گاوسی اشاره کرد که ممکن است باعث اریب شدن نتایج تخمین شود [۱۳]. همچنین طبق گفته وبستر و مکبرتنی^۳ این روش برای تعداد نمونه بیشتر از ۱۵۰ بسیار کند عمل می‌کند [۱۴].

بیشتر این روش‌ها برای برازش مدل به یک یا چند واریوگرام تجربی توسعه داده شده‌اند. با این حال در اغلب موارد به‌دست آوردن مدل واریوگرام هدف نهایی نیست؛ بلکه هدف از مدل‌سازی واریوگرام، به‌دست آوردن پارامترهای مدل برای تخمین مقادیر متغیرها در مکان‌های نمونه‌برداری نشده است [۱۵]. همچنین به طور کلی دو مشکل در رابطه با روش‌های برازش واریوگرام وجود دارد:

برای برازش مدل واریوگرام لازم است واریوگرام را به صورت دستی تحلیل و مقایسه کرد و این موضوع کاری زمان‌بر است.

روش‌های حداقل مربعات و بیشینه درست‌نمایی برای بهبود برازش واریوگرام محدود به نوع مدل واریوگرام انتخاب شده‌اند [۱۹].

با وجود اینکه نمی‌توان همواره بهترین حالت را در برازش

مدل‌سازی و بررسی پیوستگی فضایی متغیرهای نمونه‌برداری شده، همواره موضوع ضروری و چالش‌برانگیز در علوم زمین بوده است [۴-۱]. واریوگرام ابزار اساسی در مدل‌سازی فضایی و زمین‌آماري است که برای تخمین یا شبیه‌سازی متغیر پیوسته نسبت به مکان یا زمان استفاده می‌شود؛ به عبارت دیگر واریوگرام میزان همبستگی فضایی یک متغیر ناحیه‌ای را نشان می‌دهد که از رابطه ۱ به دست می‌آید [۵]:

$$\gamma(h) = E\{[Z(X_i) - Z(X_{i+h})]^2\}/2 \quad (1)$$

که در آن:

$\gamma(h)$: مقدار واریوگرام

$E\{\}$: امید ریاضی

$Z(X_i)$: مقدار متغیر در مکان X_i

$Z(X_{i+h})$: مقدار متغیر در مکان X_{i+h}

مقادیر واریوگرام تجربی تحت‌تأثیر خطای نمونه‌برداری است؛ درحالی که در مدل واریوگرام نیاز است تا نوسانات محلی نمونه‌برداری هموار شده و تابع ریاضی مانند آن به واریوگرام برازش شود. در نتیجه در اغلب موارد به جای واریوگرام تجربی، از مدل واریوگرام استفاده می‌شود [۶].

انتخاب شکل کلی مدل واریوگرام و برآورد پارامترهای آن فرآیند پیچیده‌ای است و به دانش و تجربه بالایی نیاز دارد. تعیین دستی مدل واریوگرام به این صورت است که ابتدا واریوگرام تجربی را برای چند گام محاسبه می‌کنند و سپس مدل واریوگرام با آزمون و خطا و بر اساس ظاهر بصری واریوگرام تجربی توسط کاربر برازش می‌شود. در نتیجه به یک واریوگرام تجربی می‌توان مدل‌های متفاوتی برازش کرد که انتخاب بهترین مدل یکی از چالش‌های مهم، زمان‌بر و سخت در این زمینه است [۷].

یکی از روش‌های متداول در برآورد اتوماتیک مدل واریوگرام روش حداقل مربعات وزنی^۱ است [۸]. در این روش فاصله هر نقطه از واریوگرام تجربی با مدل برازش شده به حداقل می‌رسد. این روش به دلیل حجم کم محاسبات و پیاده‌سازی ساده، یک ابزار استاندارد برای مدل‌سازی واریوگرام در بسته‌های نرم‌افزاری زمین‌آماري است. مدل‌سازی واریوگرام با روش LS به عنوان یک روش غیرمستقیم در نظر گرفته

^۲ Oliver and Webster

^۳ McBRATNEY and WEBSTER

^۱ Weighted Least Square

نقشه ویژگی تشکیل شده است یکی از محبوب‌ترین الگوریتم‌های یادگیری عمیق برای کار با داده‌های مکانی است [۱۹].

پیوند الگوریتم‌های یادگیری ماشین با علوم زمین در اکثر زیرشاخه‌ها بررسی شده است که شامل سنجش از دور در مراحل اولیه [۲۰]، ژئومورفولوژی [۲۱]، علم زمین‌جامد [۲۲]، هیدروژئوفیزیک [۲۳]، لرزه‌شناسی [۲۴]، ژئوشیمی [۲۵] و نظایر آن است. دیمیتراکوپولوس^۵ یک چارچوب کلی از هوش مصنوعی در زمین‌آمار را برای محاسبه واریوگرام تجربی و انجام کریجینگ معمولی معرفی کرد [۱۸] که پس از آن داود و ساراک^۶ از شبکه‌های عصبی برای شبیه‌سازی زمین‌آمار استفاده کردند [۱۷].

برای اولین بار لی‌چان^۷ و همکاران در سال ۱۹۸۹ مفهوم شبکه عصبی را پیشنهاد کردند [۲۶] اما به دلیل حجم زیاد محاسبات و ناپایداری در آموزش استفاده نشد. در سال ۲۰۱۲، کریژفسکی^۸ و همکاران یک مدل CNN ارائه کردند که در کاهش نرخ خطا موفق عمل کرد. مدل آن‌ها به یکی از تاثیرگذارترین کارها در زمینه بینایی ماشین تبدیل شد [۲۷] و در ادامه با اجرای آن مدل در GPU و انجام تغییراتی مانند اعمال روش Dropout [28] و Batch Normalization [29] ایجاد تغییراتی در توابع فعال‌ساز و نظایر آن است، قدرت CNN در مواجهه با مسایل پیچیده به ویژه بینایی ماشین افزایش یافت.

سورندرا پاترو^۹ و همکاران در سال ۲۰۲۲ سنگ‌ها را با استفاده از الگوریتم CNN طبقه‌بندی کردند. نقشه‌برداری زمین‌شناسی را می‌توان با استفاده از ابزارهای یادگیری ماشین خودکار کرد. الگوریتم ارائه شده می‌تواند برای بررسی سنگ‌ها و کانی‌ها در مقیاس بزرگ‌تر و همچنین برای درک ساختارهای زمین‌شناسی پیچیده استفاده شود [۳۰].

در این مقاله از شبکه عصبی کانولوشن (CNN) برای به‌دست‌آوردن پارامترهای مدل واریوگرام استفاده شده است. این روش، برازش مدل واریوگرام را به صورت خودکار انجام می‌دهد که زمان فرآیند انتخاب مدل واریوگرام را کاهش می‌دهد و دقت قابل‌قبولی دارد. رویکرد این روش شامل دو شبکه CNN است؛ اولین شبکه داده‌های پراکنده را تبدیل به تصاویر شبیه‌سازی گاوسی (SGS) می‌کند و شبکه عصبی

مدل واریوگرام تعیین کرد [۱۵]، در سال ۲۰۱۸ روشی دیگر برای بهبود مدل واریوگرام پیشنهاد شد. این روش با درون‌یابی در یک فضای چندبعدی و بالا بردن دقت برازش واریوگرام تجربی با استفاده از الگوریتم ژنتیک^۱ مدل‌های بهتری برازش می‌کند که مرتبط با واریوگرام تجربی و تخمین کریجینگ است. این روش علاوه بر اینکه عملکرد بهتری در مقایسه با مدل‌سازی واریوگرام سنتی دارد؛ روش منعطفی است و می‌تواند با کاهش زمان محاسباتی بهبود یابد. در این راستا امری^۲ سه الگوریتم برای برازش مدل پیشنهاد کرد که شامل یک نسخه اصلاح‌شده از الگوریتم GV برای برازش مدل واریوگرام با سقف مجهول، الگوریتم شبیه‌سازی تبرید برای برازش یک مدل واریوگرام با سقف محدود و الگوریتم حداقل مربعات غیرخطی برای برازش مدل Pluri-Gaussian است. هدف تمام این الگوریتم‌ها به حداقل رساندن مجموع وزنی تفاوت مجذور بین واریوگرام‌های نمونه و مدل‌سازی شده است [۱۶].

در سال ۲۰۲۲ یک روش جدید برای برازش واریوگرام با استفاده از DNN^۳ پیشنهاد شد. این روش به دلیل اینکه محدود به مدل‌های موجود در انتخاب سنتی نیست، می‌تواند بهترین برازش را برای هر توزیعی از واریوگرام تجربی به دست آورد و تجزیه و تحلیل واریوگرام رو ساده کند. از آنجا که با برازش بهترین مدل واریوگرام، واریانس کریجینگ کمینه می‌شود، پس نتایج محاسبه‌شده با این روش به نتایج درون‌یابی تئوری نزدیک‌تر است؛ در نتیجه این روش الگوریتم OK^۴ را بهینه می‌کند. مدل‌سازی واریوگرام با الگوریتم DNN به زمان محاسباتی بیشتری نیاز دارد؛ بنابراین این روش نیازمند بهبود عملکرد GPU است. با این وجود این رویکرد در مقایسه با روش‌های سنتی شبکه عصبی می‌تواند در زمان پردازش برای تجزیه و تحلیل واریوگرام صرفه‌جویی کند. این روش چون از توابع کلاسیک برای برازش استفاده نمی‌کند؛ در مدل‌های چند ساختار نیاز است تا توانایی مدل ارائه شده مورد بررسی قرار گیرد [۹].

هوش مصنوعی و الگوریتم‌های یادگیری ماشین در تجزیه و تحلیل فضایی، به‌ویژه علوم زمین و زمین‌آمار، جایگاه ویژه‌ای دارند [۱۷-۱۸]. این روش‌های مبتنی بر داده، کارایی را افزایش می‌دهند و دقت تخمین را نسبت به الگوریتم‌های آماری معمولی بهبود می‌بخشند. شبکه عصبی کانولوشنال (CNN) که از چندین

^۵ Dimitrakopoulos

^۶ Dowd and Sarac

^۷ leCun

^۸ Krizhevsky

^۹ S. Patro

^۱ Genetic Algorithms (GA)

^۲ Emery

^۳ Deep Neural Network

^۴ Ordinary Kriging

از تابع سیگموئید استفاده می‌شود. تابع سیگموئید به طور گسترده در شبکه عصبی استفاده می‌شود؛ اما استفاده آن در CNN دو اشکال عمده دارد که عبارتند از:

محوشدگی گرادیان باعث می‌شود مشتق مقادیر بسیار بزرگ یا بسیار کوچک، کوچک شود. در این حالت شبکه دیگر آموزش نمی‌بیند و پیش‌بینی‌هایش در خروجی ثابت می‌ماند.

صفر محور نبودن خروجی تابع سیگموئید که این موضوع، نرخ به‌روزرسانی وزن‌ها را کم می‌کند (رابطه ۲):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (۲)$$

ب- تابع واحد یک‌سو شده خطی^۴

این تابع مقادیر منفی را صفر و مقادیر مثبت را مقدار خودش در نظر می‌گیرد. با وجود اینکه این تابع از نظر محاسباتی بسیار کارآمد بوده و در زمینه یادگیری عمیق پرکاربرد است؛ خروجی تابع صفر محور نیست و همچنین زمانی که ورودی صفر یا نزدیک به صفر باشد، دیگر عملکردی ندارد؛ در نتیجه مقدار گرادیان صفر شده و شبکه آموزش نمی‌بیند (رابطه ۳):

$$rect(x) = \max(0, x) \quad (۳)$$

ج- تابع هزینه

تابع ضرر یا هزینه در واقع میزان خطا را در هر بار اجرای شبکه برای داده‌های آموزشی نمایش می‌دهد. این خطا در واقع تفاوت بین خروجی واقعی و مقدار پیش‌بینی شده را نشان می‌دهد. بیشتر الگوریتم‌ها بر مبنای کمینه‌کردن این تابع عمل می‌کنند؛ در واقع شبکه با کم کردن مقدار تابع هزینه، آموزش می‌بیند و بهینه می‌شود. یکی از مرسوم‌ترین توابع هزینه میانگین مربعات خطا^۵ است که به اختصار MSE گفته می‌شود. این تابع میانگین مربعات تفاوت مقدار پیش‌بینی و واقعی را طبق رابطه ۴ محاسبه می‌کند:

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n} \quad (۴)$$

که در آن:

MSE : مقدار میانگین مربعات خطا

y_i : مقدار واقعی

\hat{y}_i : مقدار پیش‌بینی شده

n : تعداد داده‌ها

کانولوشنال دوم از تصاویر SGS پارامترهای واریوگرام شامل دامنه، آزمون جهت اصلی، اثر قطعه‌ای و نسبت دامنه جهت اصلی به فرعی را به دست می‌آورد.

۲- اصول و تئوری مقاله

۲-۱- شبکه کانولوشن

CNN مجموعه‌ای از شبکه‌های عصبی است که در پردازش تصویر استفاده می‌شود. این شبکه از لایه‌های نورونی با وزن و بایاس تشکیل شده است که قابلیت یادگیری دارند. هر نورون مجموعه ورودی دریافت می‌کند و وزن‌ها و بایاس را به ورودی اعمال می‌کند و در نهایت از یک تابع غیرخطی به نام تابع فعال‌ساز استفاده می‌شود [۳۱-۳۲]. مهم‌ترین مساله در فرآیند یادگیری انتخاب یک الگوریتم بهینه‌ساز و روشی برای بهتر شدن فرآیند یادگیری است. معمولاً روش‌های یادگیری مبتنی بر گرادیان برای یک شبکه CNN انتخاب می‌شود. تابع هزینه هدف اصلی الگوریتم‌های یادگیری نظارت‌شده^۱ است. پارامترهای شبکه باید در تمام دوره‌های آموزشی^۲ به‌روزرسانی شوند و شبکه باید در تمام این دوره‌ها که نشان‌دهنده یک تکرار کامل از مقداردهی پارامترها است، به دنبال پاسخی برای بهینه‌سازی محلی باشد تا خطا را به حداقل برساند. به روزرسانی پارامترها با استفاده از روش Back Propagation انجام می‌شود. در این مرحله گرادیان هر پارامتر محاسبه می‌شود و تمامی پارامترها با توجه به تأثیری که بر خطای ایجاد شده در شبکه دارند، تغییر پیدا می‌کنند و بعد از تکرار چند دوره آموزشی بهینه‌سازی شبکه پایان می‌یابد. قسمت‌های مهم یک شبکه CNN عبارتند از:

۲-۱-۱- تابع فعال‌ساز

توابع فعال‌ساز توابعی است که یک ورودی را می‌گیرند و عملیات ریاضی روی آن انجام می‌دهند. این توابع در شبکه‌های عصبی ترکیب خطی ورودی‌ها را غیرخطی می‌کنند و می‌توانند به شبکه در یادگیری داده‌های پیچیده کمک کنند و نتایج قابل‌قبولی را در خروجی ارائه دهند. دو مورد از پرکاربردترین توابع فعال‌ساز عبارتند از:

الف- تابع سیگموئید^۳

منحنی این تابع S شکل است و مقادیر را به بازه صفر تا ۱ می‌برد. زمانی که خروجی مدل از جنس احتمال مدنظر باشد،

^۱ Supervised learning

^۲ Epoch

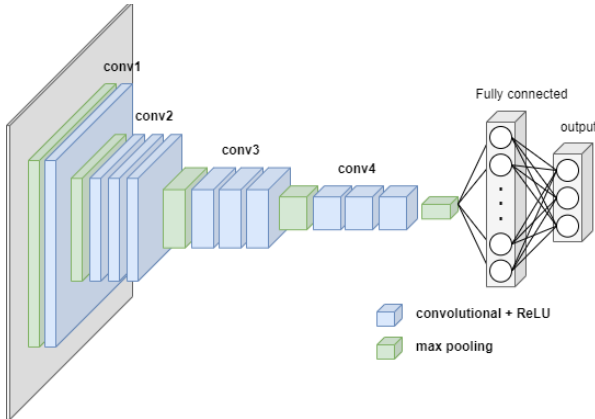
^۳ Sigmoid

^۴ ReLU / Rectified Linear Unit

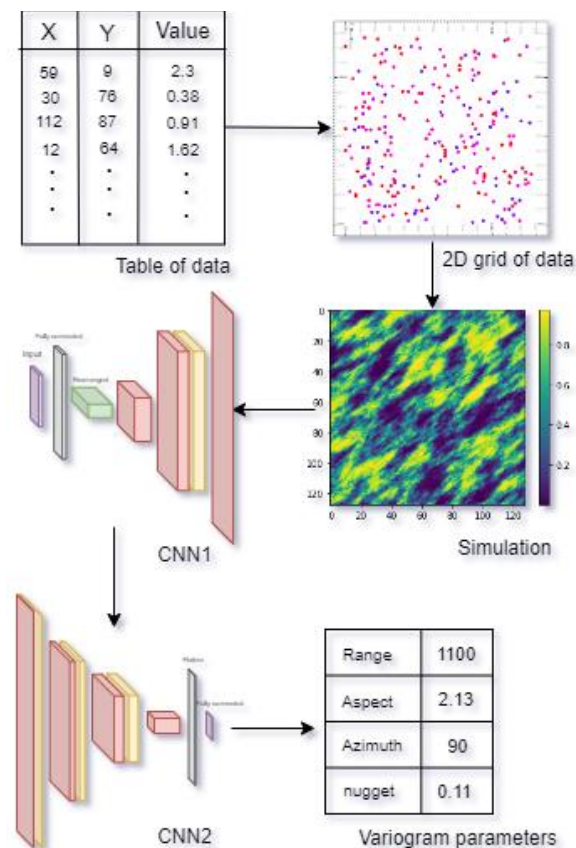
^۵ Means square error

۲-۱-۲- معماری

می‌شوند. مقادیر عیار همه نقاط با روش Nscore نرمال شده و در ادامه با استفاده از کتابخانه GeostatsPy در زبان برنامه‌نویسی پایتون ۱۰۰۰ مدل واریوگرام به صورت تصادفی با توجه به چهار پارامتر، آزیموت جهت اصلی، اثر قطعه‌ای، دامنه جهت اصلی و نسبت دامنه جهت اصلی به جهت فرعی تولید شد. این واریوگرام‌ها طبق مقادیر موجود در جدول ۱ تولید شده‌اند.



شکل ۱- معماری یک شبکه CNN



شکل ۲- فلوجارت کلی الگوریتم، با توجه به شکل داده‌های ورودی پس از شبیه‌سازی وارد شبکه CNN اول شده، خروجی شبکه CNN اول وارد شبکه CNN دوم می‌شود و پارامترها واریوگرام پیش‌بینی می‌شود

برخلاف سایر شبکه‌های عصبی که هر لایه به طور کامل به لایه بعدی متصل است، CNN از کرنل برای انتقال اطلاعات بین لایه‌ها استفاده می‌کند. CNN دارای نقشه ویژگی^۱ است که از ماتریس‌های بزرگ‌تر یا مساوی دو بعد تشکیل شده‌اند؛ درحالی‌که سایر شبکه‌های عصبی تنها از لایه‌های پنهان بین ورودی و خروجی تشکیل شده‌اند که ماتریس‌های تک‌بعدی‌اند.

الگوریتم CNN شبیه پرسپترون چندلایه عمل می‌کند؛ به همین دلیل حجم پیش‌پردازش کم می‌شود [۱۹]. بنابراین این شبکه معیارهایی را فرا می‌گیرد که در شبکه‌های قبلی به صورت دستی فراگرفته می‌شدند. CNN شامل یک لایه ورودی، یک لایه خروجی و یک لایه پنهان است که شامل چندین لایه کانولوشن، لایه ادغام^۲، لایه کاملاً متصل^۳ و نظایر آن است [۹].

لایه کانولوشن با استفاده از مجموع وزنی هر پیکسل (با اعمال وزن‌های کرنل به آن‌ها) اطلاعات را از نقشه ویژگی به نقشه ویژگی بعدی تصویر می‌کند. بعد از هر بار عمل کانولوشن که در خروجی یک نقشه ویژگی می‌دهد، یک بار فرآیند ادغام یا Upsampling انجام می‌شود که اندازه تصویر را تغییر می‌دهد. لایه ادغام ویژگی‌های موجود در نقشه ویژگی را که لایه کانولوشن تولید کرده است، خلاصه می‌کند؛ به عبارت دیگر این لایه با کوچک کردن اندازه نقشه‌های ویژگی و خلاصه‌سازی ویژگی‌های اصلی تعداد پارامترهای یادگیری میزان محاسبات شبکه را کاهش می‌دهد. لایه Upsampling عملکرد عکس لایه ادغام را دارد و اندازه تصویر را بزرگ‌تر می‌کند. معمولاً آخرین لایه شبکه CNN را لایه Fully Connected تشکیل می‌دهد. این لایه که از نوع شبکه عصبی MLP^۴ است، ویژگی‌های استخراج شده از لایه‌های کانولوشنی را طبقه‌بندی کرده و کلاس درست را برای هر ورودی شناسایی می‌کند. شکل ۱ مثالی از معماری شبکه CNN است.

۲-۲- داده‌های مصنوعی

ابتدا برای بهینه‌سازی و ارزیابی مدل یک پایگاه‌داده مصنوعی ایجاد شده است؛ در مرحله اول ۱۲۸ نقطه به صورت تصادفی طبق توزیع نرمال در یک شبکه ۱۲۸×۱۲۸ تولید شده است. هر یک از ۱۲۸ نقطه همان‌طور که در جدول موجود در شکل ۲ آورده شده است، شامل مختصات x و y و مقدار عیار در آن نقطه

^۱ Feature map^۲ Pooling layer^۳ Fully connected layer^۴ Multi-layer Perceptron

۳-۲- شبکه CNN1

این شبکه شامل یک لایه ورودی، یک لایه MLP، ۵ لایه کانولوشن، ۴ لایه Upsampling و یک لایه خروجی است.

ورودی شبکه، یک دیتابیس به صورت برداری (داده‌های استخراج شده از تصاویر شبیه‌سازی با مختصات نقاط اولیه) و خروجی شبکه یک تصویر دوبعدی (تصویر شبیه‌سازی شده) است. ابتدا لایه MLP داده‌های اولیه را گرفته و متغیرهای آن را متناسب با سایز کرنل تغییر می‌دهد؛ در مرحله بعدی لایه‌های کانولوشن وظیفه استخراج ویژگی‌های هندسی مدل را به عهده دارند و لایه‌های Upsampling نیز ابعاد نقشه را گسترش می‌دهند تا با سایز نقشه‌های اولیه برابر شوند. سایز کرنل در شبکه‌های CNN معمولاً ۳×۳، ۴×۴ یا ۵×۵ است که با آنالیز حساسیت سایز بهینه در این شبکه ۴×۴ در نظر گرفته شده است. سایر پارامترهای شبکه همچون تعداد لایه‌های کانولوشن، انتخاب تابع فعال‌ساز و نظایر آن به روش دستی و انتخابی بهینه شده است. در نهایت معماری شبکه به صورت شکل ۳ است. همانطور که در شکل ۳ مشاهده می‌شود، داده‌های ورودی پس از عبور از یک شبکه MLP، وارد لایه Reshape می‌شوند تا ابعاد متناسب با شبکه داشته باشند؛ سپس داده‌ها از چهار لایه کانولوشن همراه با Upsampling عبور می‌کنند تا بتوانند نقشه ویژگی متناسب با داده‌های اولیه را تولید کنند. در نهایت لایه آخر کانولوشن بر روی داده‌ها اعمال می‌شود و خروجی آن شکل دوبعدی شبیه‌سازی شده خواهد بود.

۴-۲- شبکه CNN2

این شبکه نیز شامل یک لایه ورودی، هشت لایه کانولوشن، پنج لایه Max pooling، یک لایه Dropout، یک لایه MLP و یک لایه خروجی است. این شبکه تصاویر دوبعدی تولید شده از CNN1 را تبدیل به پارامترهای واریوگرام (دامنه، آزیموت، نسبت دامنه جهت اصلی به فرعی، اثر قطعه‌ای) می‌کند. مراحل کار این شبکه به این صورت است که لایه‌های کانولوشن ویژگی‌های هندسی تصویر ورودی را استخراج می‌کند و خروجی آن وارد لایه Max pooling می‌شود. در این لایه ابعاد نقشه کاهش می‌یابد و ویژگی‌هایی که لایه کانولوشن تولید کرده است، خلاصه می‌شود. خروجی لایه‌های کانولوشن و Max pooling که به صورت متوالی قرار گرفته‌اند، وارد لایه Dropout شده و در نهایت در لایه MLP خروجی‌های مربوط به هر عکس شناسایی و طبقه‌بندی می‌شوند. ابعاد کرنل بهینه شده برای لایه کانولوشن ۴×۴ است که مشابه CNN1 تعیین شده است. معماری شبکه دوم در شکل ۴ نشان داده شده است.

جدول ۱- مقادیر پارامترهای واریوگرام

واحد	بازه پارامتر	پارامترهای مدل واریوگرام
درجه	صفر تا ۱۸۰	آزیموت
متر	۲۰ تا ۵۰	دامنه
بدون واحد	صفر تا ۰/۱	اثر قطعه‌ای
بدون واحد	۱ تا ۴	نسبت دامنه جهت اصلی به فرعی

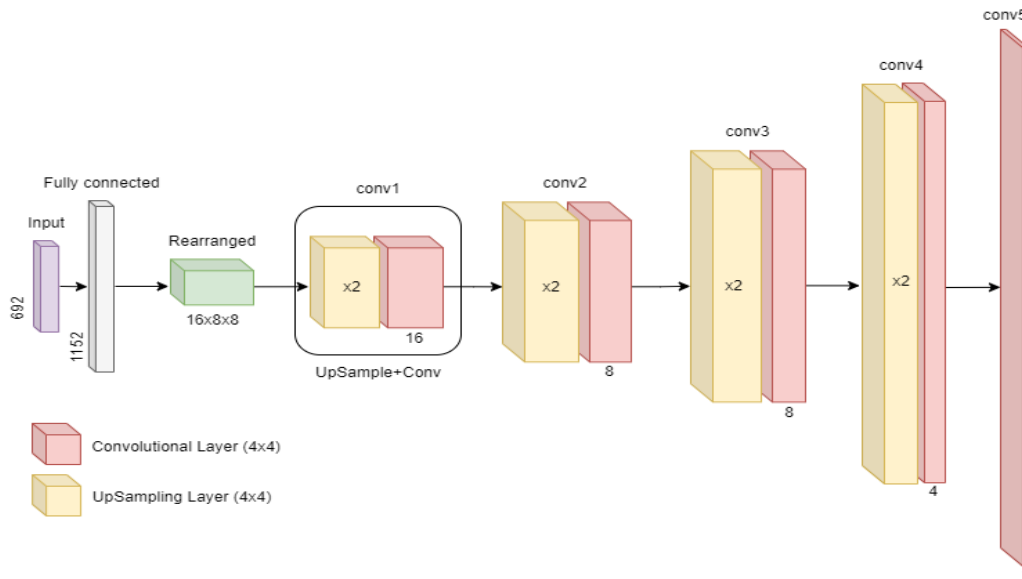
در مرحله بعد با استفاده از داده‌ها و واریوگرام‌های تولید شده، ۱۰۰۰ شبیه‌سازی با الگوریتم SGS توسط کتابخانه GeostatsPy تولید شده است؛ در مرحله چهارم، ابتدا روی مقادیر شبیه‌سازی شده به فضای اولیه بازگردانده شده و سپس با توجه به موقعیت داده‌های اولیه، از شبیه‌سازی‌های ایجاد شده داده‌های جدیدی که در همان موقعیت قرار دارند، استخراج می‌شود و این داده‌ها به عنوان ورودی در شبکه CNN1 مورد استفاده قرار گرفته‌اند. همچنین برای هر داده پارامترهای واریوگرام نیز ذخیره شده است که برچسب داده‌ها خواهند بود. این داده‌ها قبل از ورود به شبکه به سه قسمت تقسیم شده تا هر بخش برای مراحل آموزش، تست و اعتبارسنجی استفاده شود. تقسیم‌بندی داده‌ها در این مقاله به این صورت است که ۶۵ درصد داده‌ها برای آموزش، ۱۵ درصد برای اعتبارسنجی و ۲۰ درصد باقی‌مانده برای تست مدل استفاده می‌شود.

در ادامه دو شبکه CNN مورد استفاده قرار گرفته است، شبکه CNN اول داده‌های اولیه را به یک تصویر دوبعدی تبدیل می‌کند. خروجی این شبکه یک تصویر دوبعدی است که به تصویر شبیه‌سازی شده شباهت دارد. ورودی شبکه CNN دوم خروجی شبکه اول است و این شبکه وظیفه دارد تا پارامترهای مدل واریوگرام را پیش‌بینی کند. هر دو شبکه CNN با زبان برنامه‌نویسی پایتون پیاده‌سازی شده است.

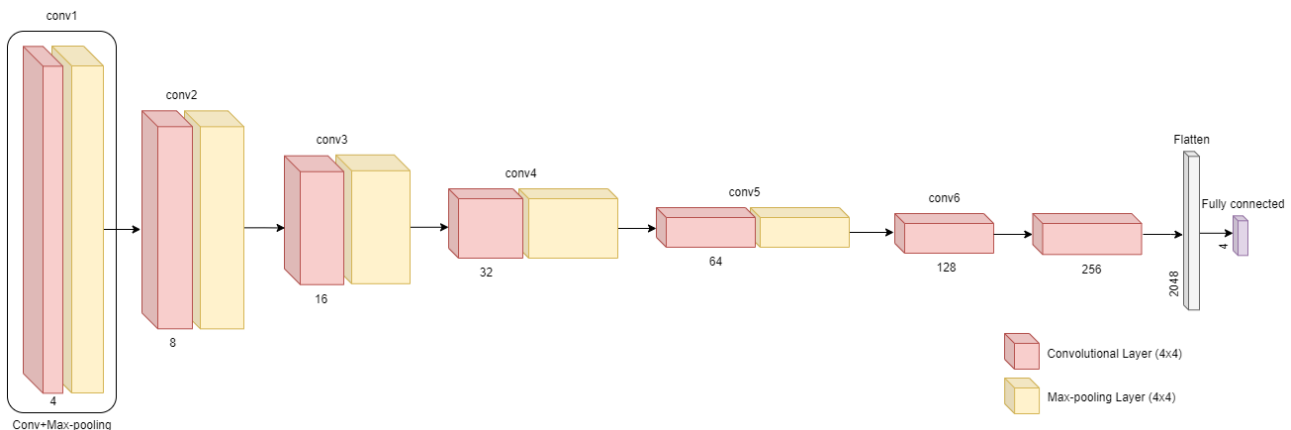
نحوه کار کلی این مدل برای پیش‌بینی پارامترهای واریوگرام در فلوچارت شکل ۲ نشان داده شده است. همان‌طور که در شکل ۲ مشخص است، ابتدا موقعیت داده‌های مورد نظر تبدیل به یک فضای ۱۲۸×۱۲۸ می‌شود، سپس با توجه به مدل‌های واریوگرام ساخته شده با جدول ۱ و داده‌های تبدیل شده یک شبیه‌سازی SGS انجام می‌پذیرد. واریوگرام هر تحقق شبیه‌سازی به عنوان برچسب آن تحقق شناخته می‌شود. سپس با استفاده از تحقق‌های ساخته شده مدل CNN اول و دوم آموزش داده می‌شود.

همچنین نرخ یادگیری برای هر دو مدل 0.001 در نظر گرفته شده است.

پروسه آموزش با به حداقل رساندن تابع هزینه که برای هر دو شبکه MSE انتخاب شده است، انجام می‌شود. تابع هزینه MSE با استفاده از روش بهینه‌سازی Adam کمینه می‌شود و



شکل ۳- معماری CNN1



شکل ۴- معماری CNN2

است CNN1 تا ۱۰۰ تکرار اول به سرعت کاهش می‌یابد و سپس تثبیت می‌شود. تابع هزینه شبکه CNN2 نیز در ۴۰۰ تکرار نشان داده شده است. نرخ کاهش و شکل نمودارها نشان از آموزش قابل قبول مدل‌ها و عدم بروز مشکل بیش برآزش در دو مدل CNN1 و CNN2 است. همچنین در مدل CNN2 همان‌طور که در نمودار دقت به تکرار نیز ترسیم شده است. همان‌طور که در شکل ۵ مشاهده می‌شود، نرخ افزایش دقت برای داده‌های آموزشی و اعتبارسنجی تقریباً برابر و همچنین دقت نهایی در این مدل 0.97 است. برای اعتبارسنجی بصری نیز نتایج مدل در شکل ۷ نشان داده شده است.

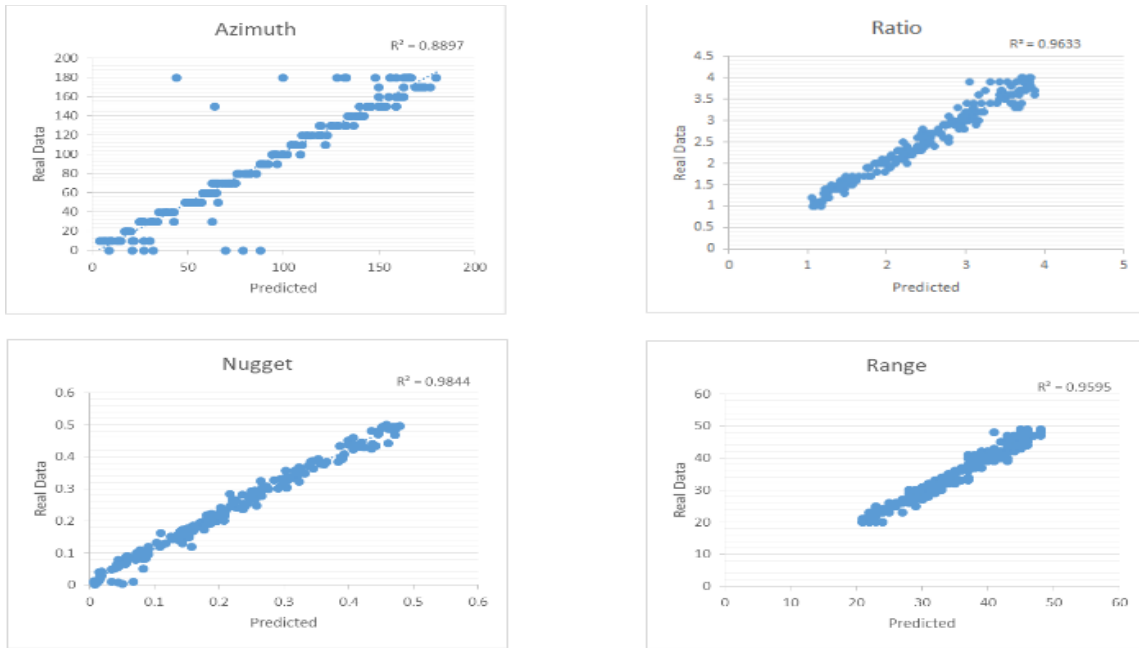
برای اعتبارسنجی دقیق‌تر نتایج به دست آمده، هر کدام از

۲-۵- اعتبارسنجی

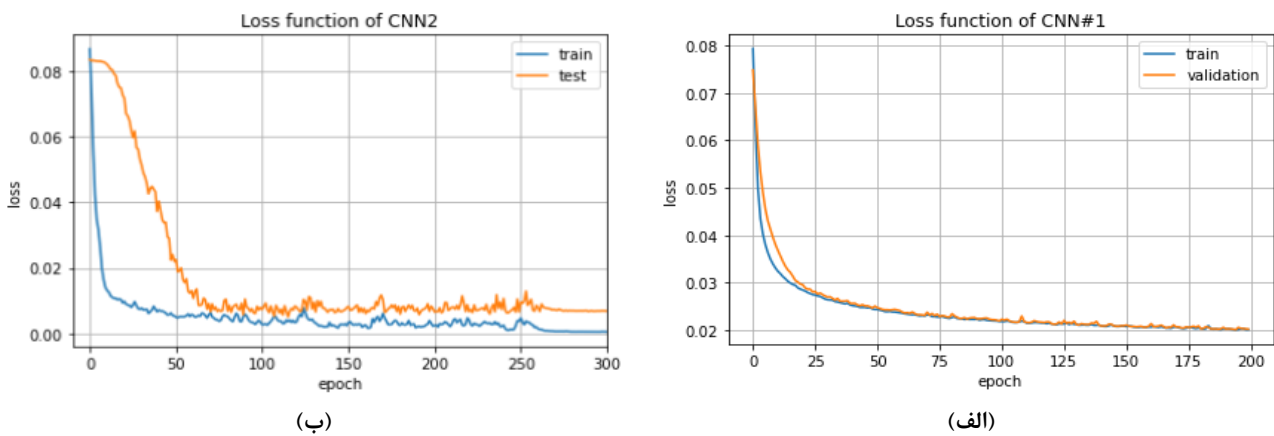
داده‌های آموزشی تنها برای آموزش مدل به کار برده شده و برای اعتبارسنجی از ۱۵ درصد داده‌های اصلی که قبل از آموزش برای اعتبارسنجی جدا شده بود، استفاده شده. مراحل یاد شده روی داده‌های اعتبارسنجی تکرار شد و با توجه به خروجی‌های به دست آمده تابع هزینه بهینه شد که هر دو شبکه در اعتبارسنجی نیز موفق عمل کرده‌اند. توابع هزینه حاصل آموزش و اعتبارسنجی مدل‌های CNN باید در هر تکرار بررسی شوند تا دقت و پایداری مدل اندازه‌گیری شده و هر گونه مشکل احتمالی در شبکه شناسایی شود. شکل ۶ نمودار تابع هزینه شبکه CNN1 و CNN2 را نشان می‌دهد. همان‌طور که مشخص

۰/۹۸ درصد بوده و نسبت دامنه جهت اصلی به فرعی، دامنه و آزمون، به ترتیب با دقت ۰/۹۶، ۰/۹۵، ۰/۸۸ پیش‌بینی شده‌اند.

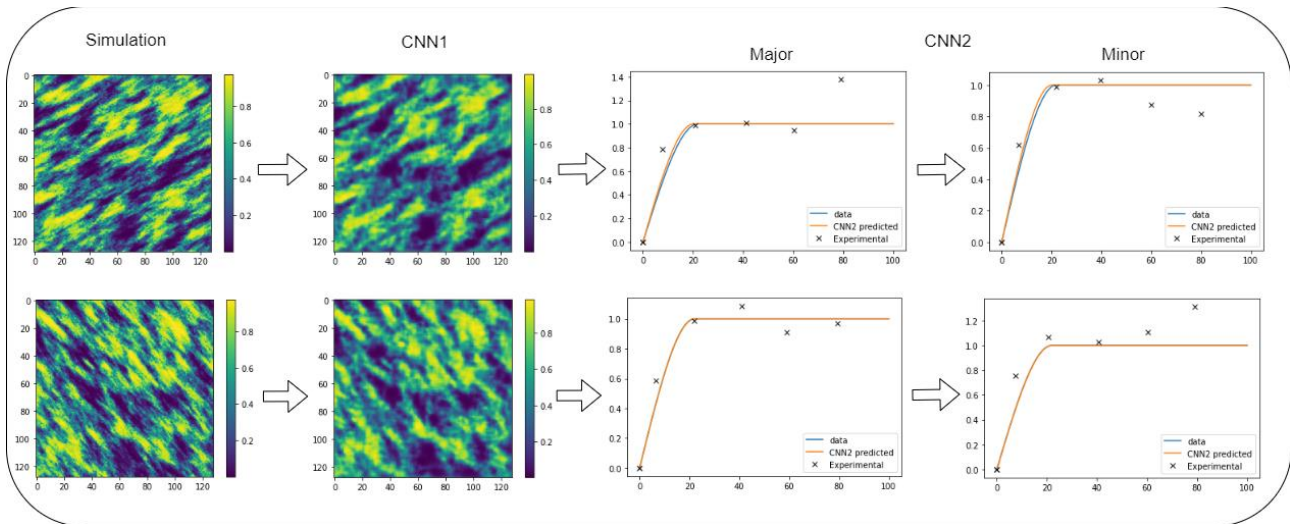
پارامترهای پیش‌بینی شده به صورت جداگانه نیز مورد بررسی قرار گرفته است. با توجه به شکل ۵ می‌توان مشاهده کرد که از بین ۴ پارامتر پیش‌بینی شده، اثر قطعه‌ای با بیشترین دقت که



شکل ۵- اعتبارسنجی پارامترهای مدل واریوگرام پیش‌بینی شده در داده‌های آزمایشی



شکل ۶- نمودار الف- تابع هزینه شبکه CNN1، ب- CNN2



شکل ۷- اعتبارسنجی بصری نتایج مدل پیشنهادی، ستون اول از سمت چپ، مربوط به داده‌های اولیه شبیه‌سازی شده است. ستون دوم خروجی مدل CNN1 را نشان می‌دهد. ستون سوم و چهارم به ترتیب نتایج خروجی‌های CNN2 است که در راستای اصلی و فرعی به صورت مدل واریوگرام ترسیم شده است.

	۶۴	۰/۰۰۰۱	۰/۰۱۱۹	۰/۹۷۰۱
۴×۴	۱۶	۰/۰۰۳۱	۰/۰۱۲۷	۰/۹۱۰۲
	۳۲	۰/۰۰۳۴	۰/۰۱۰۳	۰/۹۰۹۲
	۶۴	۰/۰۰۰۳	۰/۰۱۳۷	۰/۹۷۴۷
۵×۵	۱۶	۰/۰۰۲۷	۰/۰۱۲۷	۰/۹۴۰۹
	۳۲	۰/۰۰۱۵	۰/۰۱۱۶	۰/۹۴۷۶
	۶۴	۰/۰۰۰۹	۰/۰۱۲۲	۰/۹۵۳۹

و در جوار سنگ‌های رسوبی و آتشفشانی مربوط به دوران سوم قرار گرفته است که این سنگ‌ها به علت عملکرد تکنیکی، چین‌خورده و خرد شده‌اند. منطقه نوچون دارای توپوگرافی ناهموار و خشن و در یک ناحیه کوهستانی مرتفع واقع شده است و ارتفاع آن از سطح دریا ۲۷۰۰ تا ۲۹۵۰ متر است. این منطقه که روند شمال‌غرب-جنوب‌شرق دارد و قسمتی از بخش جنوبی زون ارومیه دختر و در قسمت انتهایی آن واقع شده است، کمربند دهج-ساردوئیه نام دارد. سنگ‌های این مجموعه بیشتر با منشا آتشفشانی است و به طور کلی این منطقه شامل سنگ‌های آذرآواری، ولکانوکلاستیک ائوسن، سنگ‌های ساب‌ولکانیک (نفوذی نیمه‌عمیق) با سن الیگومیوسن و سنگ‌های ولکانیکی نئوژن و رسوبات کواترنری است که سنگ‌های نفوذی و نیمه‌عمیق و دایک‌ها این مجموعه را در بسیاری از نقاط قطع کرده‌اند. قسمت‌های مرتفع منطقه شامل سنگ‌های ولکانیک کلاستیک و ولکانیکی است (مانند آندزیت‌های قسمت‌های شمال‌شرقی منطقه و در قسمت شمال‌غرب منطقه). سنگ‌های ساب‌ولکانیک به صورت تپه‌های بلند در منطقه قرار گرفته‌اند و در بعضی بخش‌ها به صورت تپه‌ماهور دیده می‌شوند. رسوبات کواترنری و تراست‌های آبرفتی نیز دشت جنوبی-جنوب شرقی (منطقه بین نوچون و پاریز) را به وجود آورده‌اند.

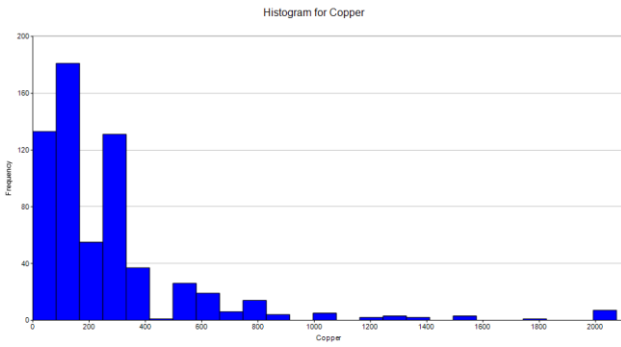
معماری شبکه با آزمایش و خطا بهینه شده است، همچنین توابع فعال‌ساز در لایه‌های کانولوشن ReLU و در لایه‌های کاملاً متصل Sigmoid در نظر گرفته شد که با توجه به مقایسه‌های انجام شده، بهترین نتیجه را نسبت به سایر توابع فعال‌ساز از خود نشان داده بودند. اما برای انتخاب سایز مناسب کرنل و BatchSize آنالیز حساسیت انجام گرفت که برای سایز کرنل سه مقدار ۳×۳، ۴×۴ و ۵×۵ و برای Batch Size مقادیر ۱۶، ۳۲ و ۶۴ در نظر گرفته شد. در جدول ۲ می‌توان نتایج مربوط به آنالیز حساسیت برای این دو پارامتر را مشاهده کرد. در این جدول مقادیر تابع هزینه برای داده‌های آموزشی و دقت شبکه نیز برای مقادیر مختلف آورده شده است. با توجه به جدول مشاهده شد که شبکه در Batch Size ۶۴ بهتر عمل کرده است. با وجود اینکه در Batch Size ۶۴ میزان خطای یادگیری و تست برای سایز کرنل ۳×۳ کمتر است؛ ولی به دلیل دقت بالای سایز کرنل ۴×۴ این سایز به عنوان بهترین سایز کرنل تعیین شد ولی با این حال با توجه به مقادیر جدول ۲ تفاوت معناداری بین این دو کرنل وجود ندارد.

۳- نتایج و بحث

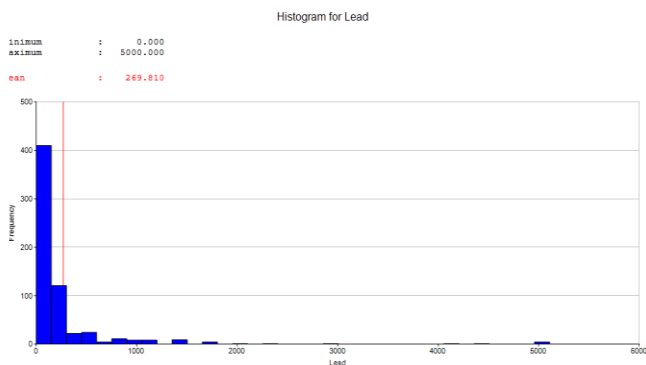
منطقه نوچون در جنوب‌شرقی ایران، در استان کرمان و ۶۵ کیلومتری شهرستان سیرجان در ۵ کیلومتری جنوب معدن مس سرچشمه و ۱۰ کیلومتری شهر پاریز قرار گرفته و دارای مختصات ۳۸۸۶۰۰ و ۳۳۱۰۵۰۰ در سیستم UTM است. همچنین این منطقه در قسمت مرکزی سلسله کوه‌های زاگرس

جدول ۲- آنالیز حساسیت پارامترهای شبکه

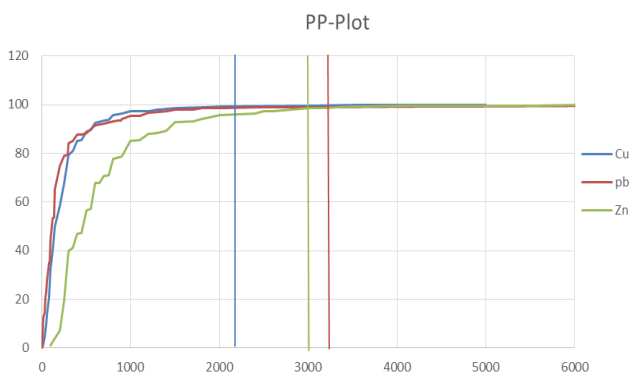
Kernel Size	Batch Size	train	validation	accuracy
۳×۳	۱۶	۰/۰۰۱۵	۰/۰۱۱۸	۰/۹۳۹۵
	۳۲	۰/۰۰۰۸	۰/۰۱۲۹	۰/۹۵۹۵



شکل ۱۰- هیستوگرام عنصر مس



شکل ۱۱- هیستوگرام داده‌های عنصر روی



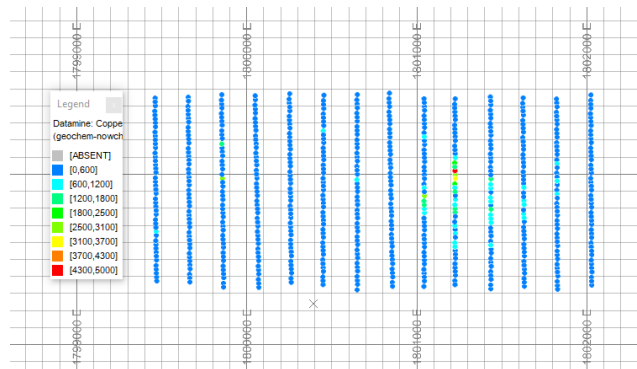
شکل ۱۲- نمودار PP-Plot سه عنصر مس، سرب و روی

در این منطقه داده‌های ژئوشیمیایی برداشت شده است. این داده‌ها شامل سه عنصر اصلی، مس، سرب و روی می‌شود و هیستوگرام هر کدام از عناصر در شکل‌های ۹، ۱۰ و ۱۱ قابل مشاهده است. با توجه به جدول ۳ که اطلاعات آماری داده‌ها را نشان می‌دهد، عیار میانگین برداشت‌شده برای عنصر مس ۲۸۰/۸۲، روی ۶۹۴/۲۰ و برای عنصر سرب ۲۹۹/۳۰ ppm گزارش شده است.

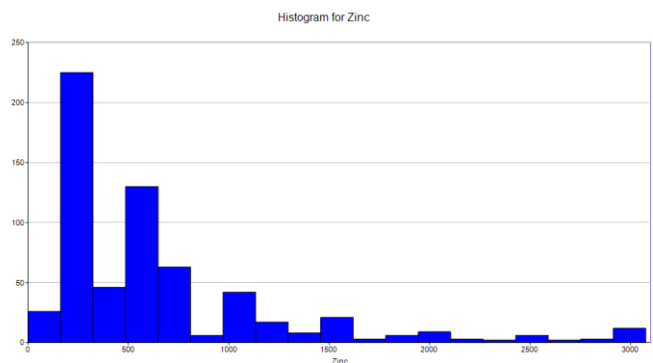
شبکه برداشت داده‌های ژئوشیمیایی همانطور که در شکل ۸ نشان داده شده است، دوعده‌ای و به شکل یک مستطیل با ابعاد ۲۵۰۰ در ۱۱۰۰ متر است که در ۱۴ نیمرخ برداشت شده است. تعداد نقاط برداشت شده ۶۳۰ نقطه بوده است که به طور متوسط در هر نیمرخ ۴۵ نمونه برداشت شده است. در نیمرخ‌های ۳ تا ۷ غلظت عناصر برداشت شده بیشتر و تمرکز آنومالی ژئوشیمیایی در نیمرخ‌های یاد شده است.

جدول ۳- آمار توصیفی داده‌ها

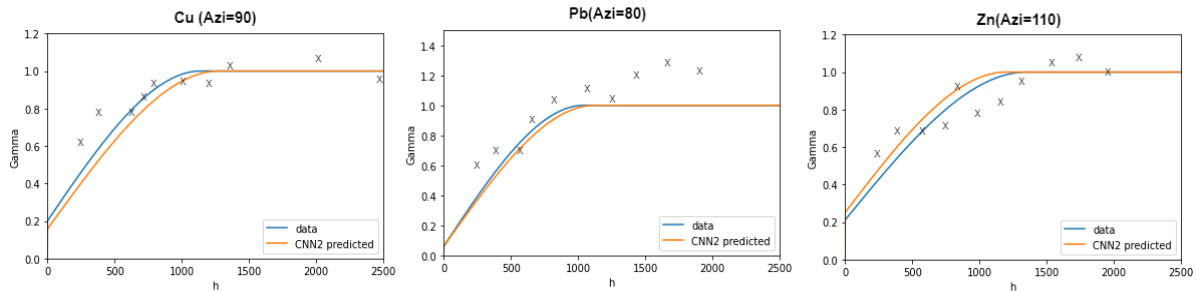
عنصر	تعداد داده	میانگین (ppm)	واریانس	انحراف معیار	میانه (ppm)
Cu	۶۳۰	۲۸۰/۸۲	۱۵۷۴۱۲/۴۰	۳۹۶/۷۵	۲۰۰
Zn	۶۳۰	۶۹۴/۲۰	۵۲۳۶۵۰/۸۰	۷۲۳/۶۳	۵۰۰
Pb	۶۳۰	۲۹۹/۳۰	۷۶۵۴۰۹/۱۰	۸۷۴/۸۷	۱۲۰



شکل ۸- شبکه برداشت داده‌ها



شکل ۹- هیستوگرام عنصر روی



شکل ۱۳- مدل واریوگرام برازش شده به صورت دستی (خط آبی) و مدل CNN (خط نارنجی) برای سه عنصر مورد مطالعه منطقه نوچون

واریوگرام		
آزیموت	صفر تا ۱۸۰	درجه
دامنه	۱۰۰۰ تا ۱۴۰۰	متر
اثر قطعه‌ای	صفر تا ۰٫۵	بدون واحد
نسبت دامنه جهت اصلی به فرعی	۲ تا ۴	بدون واحد

با توجه به جدول ۵ می‌توان نتایج پارامترهای برازش شده توسط متخصص (به صورت دستی)، مدل پیشنهاد شده در این مقاله و روش کمترین مربعات خطا مشاهده کرد. مدل پیشنهاد شده در این مقاله توانسته به خوبی آزیموت اصلی برای مدل‌سازی واریوگرام را پیش‌بینی کند که این آزیموت با آزیموت تعیین شده توسط متخصص همخوانی خوبی دارد (دقت ۰٫۹۶). در مورد پارامتر دامنه نیز شبکه CNN به خوبی توانسته است این پارامتر را پیش‌بینی کند، به طوری که در این پارامتر دقت ۰٫۹۰ است؛ اما همان‌طور که مشاهده می‌شود مدل LS برای عنصر Zn نتوانسته دامنه قابل‌قبولی را پیش‌بینی کند، علت این امر نیز همان‌طور که قبلاً گفته شده مربوط به محاسبات ریاضی و نوع همگرایی روش LS است که در برخی موارد با خطا همراه است. با توجه به اینکه در این محدوده، واریوگرافی در راستای فرعی موضوعیت ندارد. (با توجه به اینکه نمونه‌ها به صورت سیستماتیک برداشت شده‌اند و در آزیموت‌های عمود بر جهت اصلی جفت نمونه کافی یافت نمی‌شود)؛ در نتیجه مدل برازش شده به جهت فرعی اعتبار کافی ندارد و نمی‌توان به نسبت جهت اصلی به جهت فرعی به عنوان یک پارامتر قطعی استناد کرد. در نهایت اثر قطعه‌ای پیشنهاد شده با مدل، دقت قابل‌قبولی دارد؛ اما در روش LS با توجه به اینکه صرفاً از یک واریوگرام برای برازش استفاده می‌شود، مقدار اثر قطعه‌ای به طور قابل‌ملاحظه‌ای با خطا همراه است. دقت به دست آمده برای پیش‌بینی اثر قطعه‌ای در روش پیشنهاد شده ۰٫۷۵ است.

برای ورود داده‌ها به مدل CNN نیاز است تا پیش‌پردازش‌هایی انجام شود. در مرحله اول داده‌های خارج از ردیف با استفاده از روش نمودار PP-Plot مشخص شد. بر اساس نمودار شکل ۱۲، ۲ درصد از مقادیر بیشینه هر عنصر به عنوان داده خارج از ردیف در نظر گرفته شد که این مقادیر پس از شناسایی از داده‌ها کپ شدند. در مرحله بعدی تمامی داده‌ها به روش nscore نرمال شد. در نهایت برای اینکه بتوان مختصات داده‌ها را با شبکه آموزش داده شده همسان کرد، یک تبدیل مختصات بر روی داده‌ها انجام شده و سپس در شبکه منظم ۱۶۶×۴۴ با حفظ فاصله، جای‌گذاری شدند. این تبدیل مختصات واقعی داده‌ها را به یک مختصات محلی تبدیل می‌کند. مختصات پایه در این تبدیل (۰، ۰، ۰) است.

طبق پارامترهای یاد شده در جدول ۴، ۱۰۰۰ مدل واریوگرام به صورت تصادفی تولید شد؛ سپس با استفاده از داده‌های پیش‌پردازش شده و مدل‌های واریوگرام مختلف ۱۰۰۰ تحقق به روش SGS شبیه‌سازی شد. در این پروسه از ۸۰ درصد داده‌ها برای آموزش شبکه CNN1 استفاده شده و ۲۰ درصد داده برای اعتبارسنجی مدل به کار گرفته شد. در نهایت پس از اعتبارسنجی مدل نتایج این شبکه به عنوان ورودی شبکه CNN2 مورد استفاده قرار گرفت. این شبکه تصاویر ورودی را گرفته و ۴ پارامتر مدل واریوگرام را پیش‌بینی می‌کند که از ۶۰ درصد داده برای آموزش شبکه، ۲۰ درصد برای اعتبارسنجی شبکه و ۲۰ درصد از داده‌ها برای آزمایش شبکه استفاده شد. این پروسه برای هر سه عنصر به صورت جداگانه انجام گرفت. در نهایت با دادن داده‌های اولیه و واقعی به شبکه CNN1 و CNN2 در خروجی شبکه دوم پارامترهای واریوگرام برای هر عنصر پیش‌بینی شده است. مطابق شکل ۱۳ می‌توان مشاهده کرد که شبکه CNN توانسته مدل واریوگرام خوبی برای سه عنصر مورد مطالعه منطقه نوچون برازش کند.

جدول ۴- بازه پارامترهای مدل واریوگرام

پارامترهای مدل	بازه پارامتر	واحد
----------------	--------------	------

جدول ۳- اعتبارسنجی پارامترهای مدل واریوگرام برای برازش دستی، CNN و روش MSE

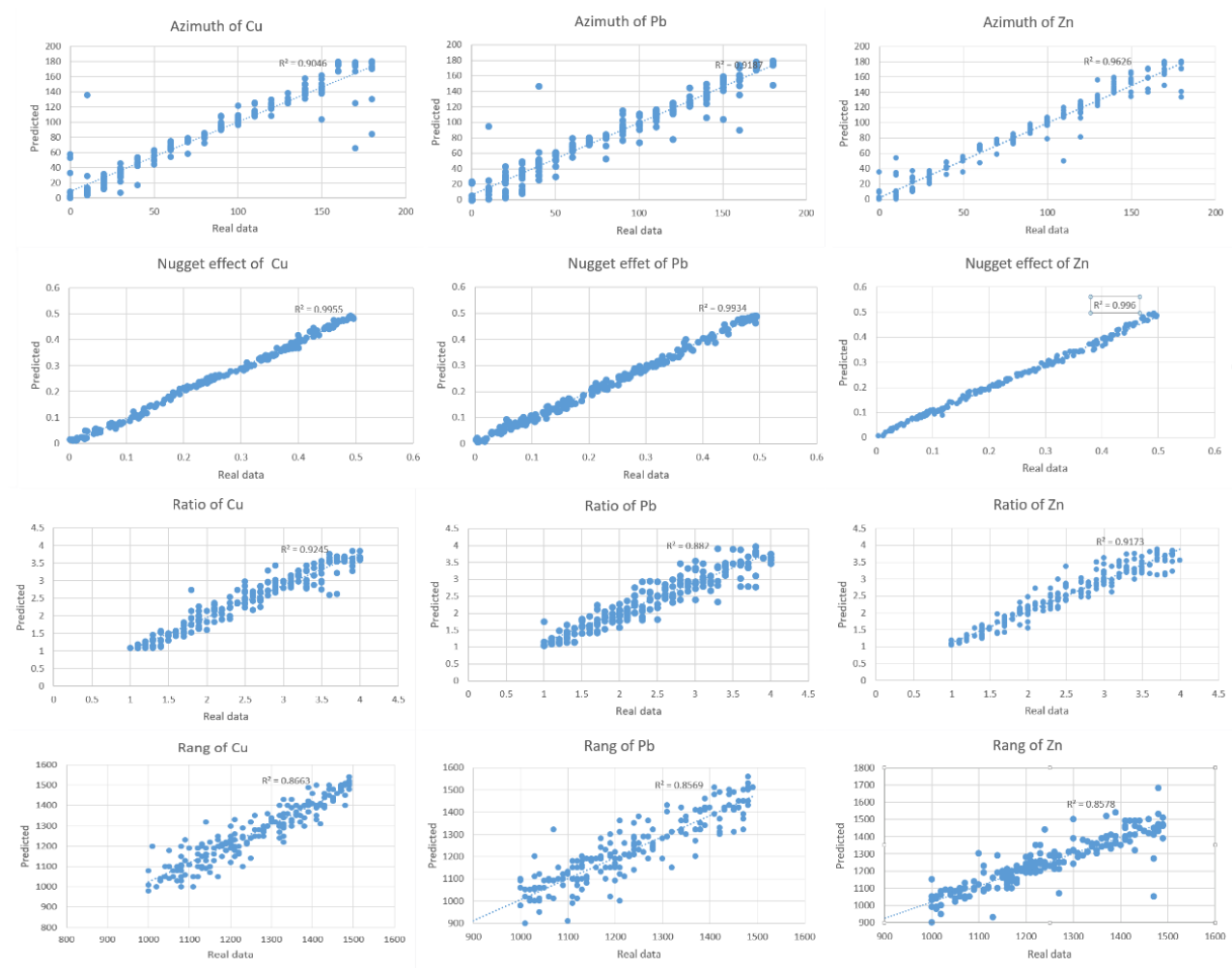
عنصر	مدل	نحوه برازش مدل	آزموت جهت اصلی	خطا (درصد)	اثر قطعه	خطا (درصد)	نسبت جهت اصلی به فرعی	خطا (درصد)	دامنه	خط (درصد)
Cu	کروی	برازش دستی	۹۰	-	۰/۲	-	۲/۸۵	-	۱۱۴۰	-
		مدل CNN	۹۳	۳/۳۳	۰/۱۵۶	۲۲	۳/۲۱	۱۲/۶۳	۱۲۷۵	۱۱/۸۴
		روش LS	۹۰	-	۰/۶	۲۰۰	-	-	۱۴۱۶	۲۴/۲۱
Pb	کروی	برازش دستی	۸۰	-	۰/۰۹	-	۲/۶	-	۱۰۴۰	-
		مدل CNN	۷۵	۶/۲۵	۰/۰۶	۳۳/۳۳	۲/۶۹	۳/۴۶	۱۱۲۰	۷/۶۹
		روش LS	۸۰	-	۰/۵۱	۸۲/۳۵	-	-	۱۰۷۹	۳/۷۵
Zn	کروی	برازش دستی	۱۱۰	-	۰/۲۱	-	۳/۲۱	-	۱۳۵۰	-
		مدل CNN	۱۱۰	۰	۰/۲۵	۱۹/۰۴	۲/۱۷	۳۲/۳۹	۱۲۰۰	۱۱/۱۱
		روش LS	۱۱۰	-	۰/۶۱	۱۹۰/۴۷	-	-	۷۰۰۰	۴۱۸/۵۲

۴- نتیجه گیری

واریوگرافی هوشمند یکی از نیازهای اساسی در مدل سازی زمین آماری است. با توجه به اینکه در یک مدل سازی زمین آماری واریوگرافی بارها انجام می پذیرد، نیاز است تا با خودکار کردن این فرآیند در زمان و هزینه صرفه جویی کرد. روش های متعددی برای برازش خودکار واریوگرام ارایه شده است اما هر کدام از این روش ها محدودیت های خاصی دارند که باعث شده تا امروز استفاده گسترده ای نداشته باشند و متخصصان ترجیح دهند تا از روش سنتی و دستی برای برازش مدل واریوگرام استفاده کنند. همچنین این روش ها توانایی پیش بینی جهت اصلی واریوگرافی را ندارند و صرفا پس از تعیین جهت اصلی مدلی بر آن برازش می کنند. روش واریوگرافی هوشمند با استفاده از شبکه های عصبی عمیق می تواند علاوه بر تعیین جهت اصلی واریوگرام پارامترهای آن را نیز پیش بینی کند.

در این مقاله یک روش مدل سازی هوشمند با یادگیری عمیق ارایه شد. این الگوریتم شامل دو شبکه CNN است.

CNN1 که یک شبکه گسترش دهنده است، داده های اولیه (نقاط مورد نظر برای واریوگرافی) و تبدیل به تصاویر دوبعدی متناظر با شبیه سازی های انجام شده می کند. وظیفه این شبکه به صورت خلاصه تهیه یک نقشه ویژگی متناسب با ارتباط فضایی میان متغیرها است. این شبکه باعث می شود تا بتوان مدل واریوگرام را ساده تر پیش بینی کرد. خروجی این شبکه وارد شبکه CNN2 می شود که یک شبکه ادغام شونده است. در این شبکه با استفاده از لایه های کانولوشن و لایه های ادغام ارتباط فضایی میان داده ها استخراج شده و تبدیل به پارامترهای واریوگرام می شود. با توجه به نتایج به دست آمده در این مقاله طبق شکل ۱۴ می توان گفت دقت شبکه پیشنهادی برای تعیین پارامترهای واریوگرافی به ترتیب ۰/۹۹ درصد برای اثر قطعه ای، ۰/۹۰ درصد برای نسبت دامنه جهت اصلی به فرعی، ۰/۸۶ درصد برای دامنه و ۰/۹۲ درصد برای آزموت است. دقت شبکه پیشنهادی به وسیله داده های آزمایشی و آموزشی به دست آمده است، همچنین برای بررسی اختلاف مدل پیشنهاد داده شده، با تصمیم متخصص خطای نسبی میان این دو روش بیان شده که در جدول ۵ به آن اشاره شده است.



شکل ۱۴- اعتبارسنجی پارامترهای مدل واریوگرام پیش‌بینی شده برای داده‌های منطقه نوچون

modeling. In: Deutsch CV”.

4. C. V. Deutsch and A. G. Journel, “GSLIB: geostatistical software library and user’s guide. Second edition,” *GSLIB geostatistical Softw. Libr. user’s Guid. Second Ed.*, 1998.
5. M. E. Rossi and C. V. Deutsch, *Mineral resource estimation*. 2014. doi: 10.1007/978-1-4020-5717-5.
6. R. A. Olea, “Fundamentals of Semivariogram Estimation, Modeling, and Usage,” *Stoch. Model. Geostatistics*, pp. 27–35, 1994.
7. X. Jian, R. A. Olea, and Y. S. Yu, “Semivariogram modeling by weighted least squares,” *Comput. Geosci.*, vol. 22, no. 4, pp. 387–397, 1996, doi: 10.1016/0098-3004(95)00095-X.
8. C. A. Gotway, “Fitting semivariogram models by weighted least squares,” *Comput. Geosci.*, vol. 17, no. 1, pp. 171–172, 1991, doi: 10.1016/0098-3004(91)90085-R.
9. Y. Li, Z. Baorong, X. Xiaohong, and L. Zijun, “Application of a semivariogram based on a deep neural network to Ordinary Kriging interpolation of elevation data,” *PLoS One*, vol. 17, no. 4 April, pp. 1–12, 2022, doi: 10.1371/journal.pone.0266942.
10. M. A. Oliver and R. Webster, “A tutorial guide to geostatistics: Computing and modelling variograms and kriging,” *Catena*, vol. 113, pp. 56–

از محدودیت‌های این روش می‌توان به دوبعدی بودن آن اشاره کرد. همچنین در الگوریتم پیشنهادی، مدل واریوگرام برازش شده تک ساختار است که می‌توان با اضافه کردن پارامترهای بیشتر به مدل و افزایش عمق کرنل شبکه، از مدل‌های چندساختار نیز استفاده کرد. همچنین با توجه به اینکه برای هر مدل واریوگرام نیاز به شبیه‌سازی اولیه است، این پروسه می‌تواند زمان بر باشد که در صورت اجرای برنامه بر روی GPU می‌توان در زمان اجرای برنامه صرفه‌جویی کرد.

منابع

1. M. David and R. A. Blais, “Geostatistical Ore Reserve Estimation,” pp. 27–33, 1973, doi: 10.2307/2286639.
2. DSS, J. L. Jensen, L. W. Lake, P. W. M. Corbett, and D. J. Goggin, “Statistics for Petroleum Engineers and Geoscientists,” *J. Am. Stat. Assoc.*, vol. 93, no. 442, p. 844, 1998, doi: 10.2307/2670149.
3. D. C. Larrondo PF, Neufeld CT, “VARFIT: A program for semi-automatic semivariogram

- Handwritten Zip Code Recognition,” *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989, doi: 10.1162/neco.1989.1.4.541.
27. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Communications of the ACM*, 2017, vol. 60, no. 6, pp. 84–90. doi: 10.1145/3065386.
 28. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.
 29. S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” *32nd Int. Conf. Mach. Learn. ICML 2015*, vol. 1, pp. 448–456, 2015.
 30. S. Patro, D. C. Jhariya, M. Sahu, P. Dewangan, and P. Y. Dhekne, “Igneous rock classification using Convolutional neural networks (CNN),” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1032, no. 1, 2022, doi: 10.1088/1755-1315/1032/1/012045.
 31. A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.
 32. Y. LeCun, K. Kavukcuoglu, and C. Farabet, “Convolutional networks and applications in vision,” *ISCAS 2010 - 2010 IEEE Int. Symp. Circuits Syst. Nano-Bio Circuit Fabr. Syst.*, pp. 253–256, 2010, doi: 10.1109/ISCAS.2010.5537907.
 - 69, 2014, doi: 10.1016/j.catena.2013.09.006.
 11. K. V. Mardia and R. J. Marshall, “Maximum likelihood estimation of models for residual covariance in spatial regression,” *Biometrika*, vol. 71, no. 1, pp. 135–146, 1984, doi: 10.1093/biomet/71.1.135.
 12. E. Pardo-Igúzquiza, K. V. Mardia, and M. Chica-Olmo, “MLMATERN: A computer program for maximum likelihood inference with the spatial Matérn covariance model,” *Comput. Geosci.*, vol. 35, no. 6, pp. 1139–1150, 2009, doi: 10.1016/j.cageo.2008.09.009.
 13. N. Cressie, “Fitting variogram models by weighted least squares”.
 14. A. B. McBRATNEY and R. WEBSTER, “Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates,” *J. Soil Sci.*, vol. 37, no. 4, pp. 617–639, 1986, doi: 10.1111/j.1365-2389.1986.tb00392.x.
 15. R. A. Olea, *Geostatistics for Natural Resources Evaluation by Pierre Goovaerts*, vol. 31, no. 3. 1999.
 16. X. Emery, “Iterative algorithms for fitting a linear model of coregionalization,” *Comput. Geosci.*, vol. 36, no. 9, pp. 1150–1160, 2010, doi: 10.1016/j.cageo.2009.10.007.
 17. A. neural network approach to geostatistical Simulation, “Dowd, PA; Sarac, C”.
 18. R. Dimitrakopoulos, “Artificially intelligent geostatistics: a framework accommodating qualitative”.
 19. H. Jo and M. J. Pyrcz, “Automatic Semivariogram Modeling by Convolutional Neural Network,” *Math. Geosci.*, vol. 54, no. 1, pp. 177–205, 2022, doi: 10.1007/s11004-021-09962-w.
 20. D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, “Machine learning in geosciences and remote sensing,” *Geosci. Front.*, vol. 7, no. 1, pp. 3–10, 2016, doi: 10.1016/j.gsf.2015.07.003.
 21. A. Valentine and L. Kalnins, “An introduction to learning algorithms and potential applications in geomorphometry and Earth surface dynamics,” *Earth Surf. Dyn.*, vol. 4, no. 2, pp. 445–460, 2016, doi: 10.5194/esurf-4-445-2016.
 22. K. J. Bergen, P. A. Johnson, M. V. De Hoop, and G. C. Beroza, “Machine learning for data-driven discovery in solid Earth geoscience,” *Science (80-.)*, vol. 363, no. 6433, 2019, doi: 10.1126/science.aau0323.
 23. C. Shen, “A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists,” *Water Resour. Res.*, vol. 54, no. 11, pp. 8558–8593, 2018, doi: 10.1029/2018WR022643.
 24. Q. Kong, D. T. Trugman, Z. E. Ross, M. J. Bianco, B. J. Meade, and P. Gerstoft, “Machine learning in seismology: Turning data into insights,” *Seismol. Res. Lett.*, vol. 90, no. 1, pp. 3–14, 2019, doi: 10.1785/0220180259.
 25. R. Zuo, Y. Xiong, J. Wang, and E. J. M. Carranza, “Deep learning and its application in geochemical mapping,” *Earth-Science Rev.*, vol. 192, pp. 1–14, 2019, doi: 10.1016/j.earscirev.2019.02.023.
 26. Y. LeCun *et al.*, “Backpropagation Applied to