Detection of damage status of buildings in the 2016 Kumamoto Earthquake using deep learning with on-site photos

W. Liu[(1)], K. Tozawa[(2)], Y. Maruyama[(3)], F. Yamazaki[(4)], M. Matsuoka[(5)], K. Horie[(6)]

[(1)] *Assistant Professor, Graduate School of Engineering, Chiba University, wen.liu@chiba-u.jp*
[(2)] *Undergraduate Student, Department of Urban Environment Systems, Chiba University, k_tozawa0129@chiba-u.jp*
[(3)] *Professor, Graduate School of Engineering, Chiba University, ymaruyam@faculty.chiba-u.jp*
[(4)] *Research Fellow, National Research Institute for Earth Science and Disaster Resilience, fumio.yamazaki@bosai.go.jp*
[(5)] *Professor, School of Environment and Society, Tokyo Institute of Technology, matsuoka.m.ab@m.titech.ac.jp*
[(6)] *Fellow, MS&AD InterRisk Research & Consulting, kei.horie@ms-ad-hd.com; ken.kawabe@ms-ad-hd.com*

Abstract

A series of earthquakes hit Kumamoto Prefecture in Kyushu Island, Japan, in April 2016. Due to the strong shaking, 8,667 houses were collapsed, and more than 200 thousand buildings were damaged. The government of Kumamoto Prefecture issued a total of 203,882 disaster victim certificates until May 11, 2017. In disaster victim certificates, the damage statuses of target buildings were classified into five levels: no damage, minor damage, moderate– damage, moderate+ damage, and major damage, according to field investigation. For each building, at least one photo was taken by the investigator as proof data. The investigation was carried out from April to October 2016, which cost human efforts and time. It also caused the delay of the issuance of disaster victim certificates, which were essential when applying for various disaster assistance services. The objective of this research is to propose an effective and objective procedure to improve the discrimination of building damage and to shorten the issuance time of disaster victim certificates. As the first step, six convolutional neural network (CNN) models were tested to identify major-damage buildings from the external photos of buildings.

Two data sets were created manually, including 7,873 on-site photos of building appearances taken in Kumamoto City, Uki City, and Mashiki Town. An external photo should include the building's roof, walls, and foundation clearly without other objects, e.g. humans or vehicles. A mirror inversion and rotation processing were applied to increase the number of photos for training. Finally, the number of training data was 4,772 photos, whereas that of the test data was 1,194 photos. The ratio of the training data and the test data was 8:2. Those photos were labeled into two classes, major damage and others. The first data set consisted of the buildings in each damage level, whereas the second data set included the buildings in major damage, minor damage, and no damage levels. These two data sets were fed to six CNN models. Four models were modified from the VGG16 model with different dropout rates, whereas the other two models were based on the ResNet model with different numbers of convolutional layers. The number of iterations was set to be 30,000. Every 100 iterations, the accuracy and the loss were evaluated using the test data. Comparing with the accuracy and the loss obtained at the 30,000th iteration, the modified VGG16 model without the dropout layer using the second data set obtained the best result. The accuracy reached 77.8% and the loss was 1.28.

Keywords: convolutional neural network, the 2016 Kumamoto earthquake, collapsed building, disaster victim certificate

1. Introduction

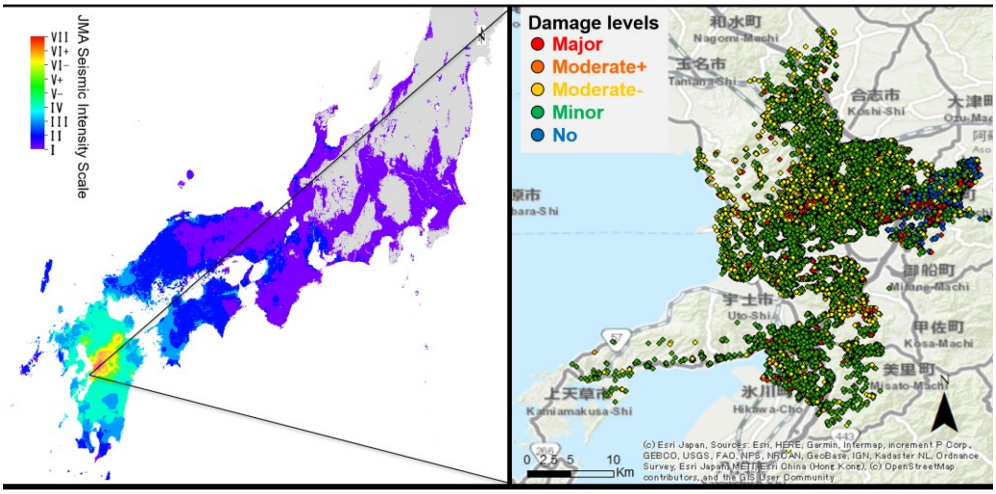
An Mw 6.2 earthquake hit the Kumamoto Prefecture in Kyushu Island, Japan, at 21:26 (JST) on April 14, 2016. The epicenter was located in the Hinagu fault with a shallow depth. Twenty-eight hours later, another Mw 7.0 earthquake occurred in the Futagawa fault at 01:25 (JST) on April 16, close to the Hinagu fault. The first event was called a "foreshock" and the second one the "mainshock". Strong shaking, level 7 on the Japan Meteorological Agency (JMA) seismic intensity scale was observed in Mashiki Town in both the foreshock and mainshock. The JMA seismic intensity of the mainshock is shown in **Fig. 1**(a), cited from QuiQake [1]. A total of 273 people was killed by related causes in the earthquake sequence. 8,667 houses were collapsed, and more than 200 thousand buildings were damaged due to this event [2].

After a disaster occurs, a local government in Japan issues disaster victim certificates to damaged buildings. The damaged buildings would be classified into five levels based on field investigation with at least one on-site photo. The disaster victim certificate is essential when applying for various disaster assistance services. Thus, it is important to issue disaster victim certificates quickly and accurately. However, according to the statues of past disasters, the issuance of the disaster victim certificates took a long time when the affected area was large. In the event of the Kumamoto earthquake, a total of 203,882 disaster victim certificates were issued in Kumamoto Prefecture until May 11, 2017 [3]. However, the certificates issued successfully in the first month were less than 30% of the requests. The delay of the certificate issuance was a severe problem for the life recovery of victims. The investigation was carried out using a paper format from April to October 2016, which cost human efforts and time. Besides, the damage levels was assessed through the field survey by human investigators. Therefore, its results included subjective errors. An effective and objective discrimination procedure of damage levels would help to shorten the issuance time of the disaster victim certificates.

In recent researches, Convolutional Neural Networks (CNN) outperform traditional methods on the damage assessment. CNN is one of the deep learning algorithms, having good potential in image recognitions. Vetrivel et al. [4] developed a framework for damage classification using CNN features and 3D point cloud features. Duarte et al. [5] proposed three multi-resolution CNN feature fusion approaches and applied them to several satellites and airborne image samples. Nex et al. [6] tested the performance of an advanced CNN on satellite, airborne and UAV images. On-site photos were also utilized to determine building damages by CNN algorithms [7-8]. In this study, we applied CNN architectures to the on-site photos taken during the field investigations after the Kumamoto earthquake. The objective of this study is to improve the efficiency of damage assessment and to reduce the issuance time of the disaster victim certificates. As the first step, the discrimination of major damaged buildings from external images was conduct and the result was examined.

2. Date sets and CNN models

To issue a disaster victim certificate, professional investigators would conduct a field survey for a target building. The judgment of the damage levels is according to the flowchart defined by the Cabinet Office of Japanese Government [9]. The building would be assigned to five classes: major damage, moderate+ damage, moderate- damage, minor damage, and no damage. The discrimination is carried out according to the outside appearance firstly. If the building obviously collapses, it would be classified as "major damage". If the building is still standing, the inclination of walls or pillars is measured next. When the inclination is over 9°, the building is classified as "major damage", otherwise the partial damage percentage would be calculated. If more than 75% of the building foundation is damaged or the total monetary loss is over 50%, it would be classified as "major damage". When the total monetary loss is between 40% to 50%, it would be classified as "moderate+ damage". When the total loss is between 20% to 40%, it would be classified as "moderate- damage". When the total loss is less than 20%, it would be classified as "minor damage". During the field survey, the investigators took several on-site photos as the proof of damage classification. In this study, we used those on-site photos to perform the first judgment of the outside appearance automatedly.



(a) (b)

Fig. 1 – (a) JMA seismic intensity in the mainshock of the 2016 Kuamamoto earthquake by QuiQuake [1]. (b) Distribution map of building damage in Kumamoto City, Mashiki Town and Uki City, Kumamoto Prefecture, Japan, according to the disaster victim certificates.

Table 1 –The number of the buildings in each damage level

Jurisdiction	Damage level					Total
	Major	Moderate+	Moderate-	Minor	No	
Kumamoto City	3,759	5,208	21,554	34,659	3,240	68,420
Uki City	522	325	1,655	5,727	18	8,247
Mashiki Town	4,883	941	2,985	7,747	2,941	19,497

The data was collected from the field investigations by Kumamoto City, Mashiki Town, and Uki City, Kumamoto Prefecture. The number of investigated buildings in each damage level is shown in **Table 1**. This study consideres only the wooden houses. 20,118 wooden houses in Kumamoto City and 7,367 wooden houses in Mashiki Town were selected as targets, including 115,738 photos of Kumamoto City and 51,695 photos of Mashiki Town.

Then the external photos of target buildings were picked out manually. An external photo should consist of building's roof, walls, foundation clearly without other objects, e.g. humans or vehicles. As a result, 6,394 external photos of Kumamoto City and 1,440 external photos of Mashiki Town were selected. The example of the on-site photos in each damage level is shown in **Fig. 2**. The number of external photos in each damage level is shown in **Table 2**. The photos in the classes of major damage and no damage were less than the other classes. Then 39 external photos of major damaged buildings in Uki City, 80 external photos of no damage buildings from the Internet were added. An adjustment of the number of photos in each damage level was carried out. Then these photos were divided randomly into two groups, training data and test data with the ratio of 1:1. For the training data, mirror inversion and rotation (±90°) processing were applied to increase the number of photos. All the photos were resized to 64 ×64 pixels and transformed into grayscale. Finally, the number of training data was 4,772, whereas that of the test data was 1,194. The ratio of the training data and the test data was 8:2.

Two data sets (I and II) were created using the prepared photos. Their breakdown is shown in **Table 3**. Each data set includes two classes. Class 1 is the major damage, and Class 2 is the other damage levels. In the data set I, Class 2 included the external photos in each damage level except the major damage. Since the

On-site external photos					
Damage level	Major	Moderate+	Moderate-	Minor	No

Fig. 2 – Example of the on-site photos of wooden house appearence in each damage level

Table 2 – The number of external photos in each damage level

Jurisdiction	Damage level					Total
	Major	Moderate+	Moderate-	Minor	No	
Kumamoto City	746	355	1,010	3,217	1,066	6,394
Mashiki Town	441	175	289	455	80	1,440

Table 3 – Breakdown of the two data sets. The number in () is the photos used in training process.

(a) Data set I

Class	1	2			
Damage level	Major	Moderate+	Moderate-	Minor	No
Number of photos	2966	775 (620)	775 (620)	775 (620)	675 (540)
Total	2,966 (2,372)	3,000 (2,400)			

(b) Date set II

Class	1	2			
Damage level	Major	Moderate+	Moderate-	Minor	No
Number of photos	2966	0	0	1,850 (1,480)	1,150 (920)
Total	2,966 (2,372)	3,000 (2,400)			

number of photos without damage was limited, that of no damage was a little less than those of other damage levels. The total numbers of Classes 1 and 2 were adjusted intentionally. In the data set II, Class 2 included only the photos in the minor and no damage levels. The number of photos in the minor and no damage levels was twice as large as that in the data set I. The total number of photos in Class 2 was the same as that of data set I.

A CNN consists of an input, an output layer, and multiple convolutional layers. In this study, two types of architectures were used to classify the damage levels of buildings. The first architecture is a modified VGG16 model [10], as shown in **Fig. 3**. It consists of 13 convolutional layers, five max-pooling layers, two full-connected layers. Since the input image size was 64 ×64 pixels, smaller than the original model (224 ×224 pixels), one full-connected layer was removed. In addition, the batch normalization was introduced to reduce the dependence on initial values.

Dropout regularization is an effective way to regularize a deep neural network and avoid overfitting [11]. We added 13 dropout layers continuously after each convolutional layer to obtain a new model. The dropout rate was changed as 0.2, 0.5 and 0.9. This model was also applied to the prepared data.

The second architecture is the ResNet model. The ResNet model is proposed by He et al. [12] including features: heavy batch normalization and skip connections. In this study, we applied two different models based

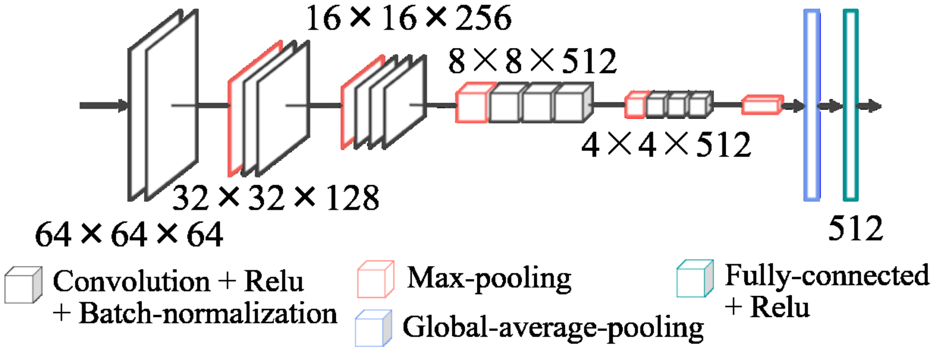


Fig. 3 – The proposed architecture of CNN based on the VGG16 model

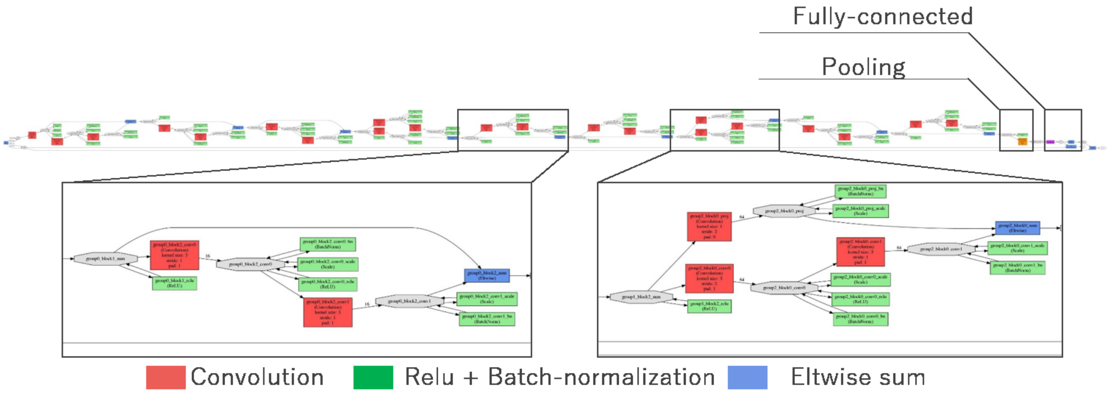


Fig. 4 – The ResNet model with 20 weight layers

Table 4 – Six CNN models employed in this study

CNN Models	Number of layers				Dropout rate
	Convolution	Pooling	Full-connected	Dropout	
1	13	5	2	0	-
2				13	0.2
3					0.5
4					0.9
5	43	1	1	0	-
6	19	1	1	0	-

on ResNet. One has 20 weight layers and the other has 44 weight layers. The architecture of 20 weight-layers model is shown in **Fig. 4**.

Six models were created to identify major damage buildings. The comparison of the six models is shown in **Table 4**. Models 1 to 4 are the modified VGG16 models with different dropout rates. Models 5 and 6 are developed from the ResNet models with different depths of convolution layers.

3. Results and discussions

The models are implemented using Caffe [13] operating on Ubuntu18.04.2 with an NVIDIA GeForce GTX 980 GPU. The optimizer used during training is Nesterove Momentum [14]. Different from the normal momentum optimization, the gradient is not computed from the current position but slightly ahead in the direction of the momentum from an intermediate position. The equations of the update rule are shown in Eqs. (1)-(2), where γ is a random constant parameter between 0 and 1, η is the learning rate. Comparing with the other optimizers, e.g. Adam or SGD, this optimization was the fastest and thus adopted in this study.

$$x_{t+1} = x_t + v_t - \eta \nabla f(x_t + v_t) \tag{1}$$

$$v_t = \gamma(x_t - x_{t-1}) \tag{2}$$

The initial learning rate was set as 0.1, and it would be multiplied by 0.1 every 10,000 iterations. The training process was conducted in 30,000 iterations. Every 100 iterations, the accuracy and the loss were evaluated. The accuracy was obtained by dividing the number of correct predictions from the total number of predictions. The cross-entropy loss was adopted as the loss function. Its value increases as the predicted probability diverges from the actual label.

The two prepared data sets including 5,966 external photos of the building appearance were fed to the six models for training and verification. The accuracy and the loss calculated using the test data at the 30,000 iterations are shown in **Table 5**. The accuracy for Data set I, including the photos in each damage level, was between 50% to 69%. The accuracy for Data set II, including the photos of major damage, minor or no damage, was between 50% to 78%. The best accuracy was obtained by model 1 using Data set II, which reached 77.8%. The evaluation of this combination in terms of accuracy and loss is shown in **Fig. 5**.

Comparing the same model fed by the different data sets, the results using Data set II showed better accuracy than that using Data set I. Several moderate+ damaged buildings were misclassified to the major damage class. It caused a decrease in the accuracy of Data set I. When two neighboring buildings in the different damage levels are reflected in the same photo, the proposed models are difficult to identify the target buildings successfully. The loss for Model 4 with 0.9 dropout rate was the lowest in **Table 5**, however, the accuracy for the training was level off from 50%. It means the learning process did not work well. As the dropout rate increases, fewer features were counted in the training. The deviation between the training loss and the test loss was observed in all the models. Generally, the deviation is caused by overfitting. According to the results of Models 2 to 5, the introduction of dropout layers did not increase the performance. Thus, the deviation that occurred in this study was caused by a few numbers of photos. Comparing the results from Models 5 and 6, the addition of convolutional layers did not improve the performance.

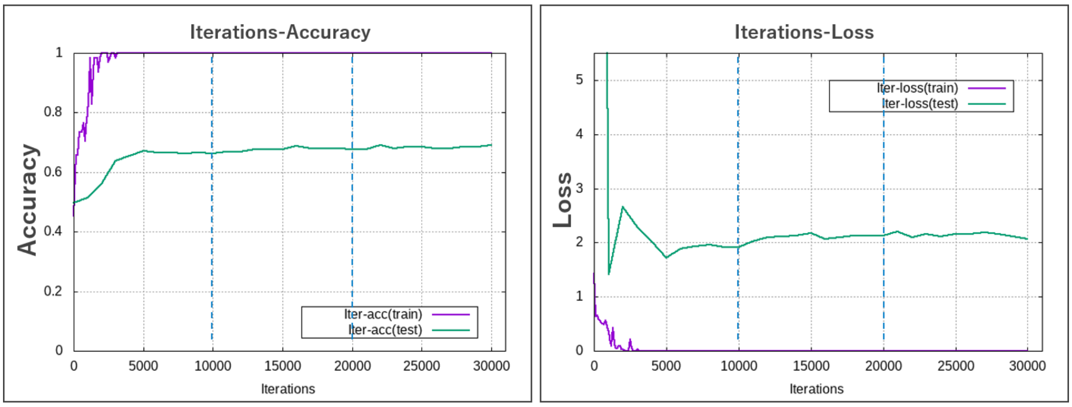
Table 5 – Evalution of the results of the used models

(a) Date set I

Models	1	2	3	4	5	6
Accuracy [%]	69.0	66.3	49.7	50.3	60.3	63.3
Loss	2.07	2.85	1.84	0.69	1.90	2.21

(b) Date set II

Models	1	2	3	4	5	6
Accuracy [%]	**77.8**	68.8	53.9	50.7	67.7	68.7
Loss	**1.28**	2.19	4.96	0.70	1.60	2.00



(a) (b)

Fig. 5 – Evaluation of the model 1 using Data set I in terms of (a) the accuracy, and (b) the loss

4. Conclusions

In this study, six CNN models were fed to the data sets of on-site photos taken during the field survey after the 2016 Kumamoto earthquake. Our goal is to propose an automated procedure for the discrimination of building damage levels and to shorten the issuance of the disaster victim certificates. As the first step, this study tried to detect the buildings with major damage.

7,873 external photos of building appearances taken in Kumamoto City, Uki City, and Mashiki Town and 80 photos of no damage buildings from the Internet were selected manually to create two data sets. Each data set included 4,772 photos for training and 1,194 for testing. In the six CNN models, four of them were modified from the VGG16 model with different dropout rates, whereas the other two models were based on the ResNet model with different numbers of convolutional layers. Comparing those testing results at the 30,000th iterations, the modified VGG16 model without dropout layers showed a better result than the other models. The models using the data set II that only included the photos of major damage, minor and no damage levels, showed better results than the models using the data set I that included the photos of all damage levels. The best accuracy reached 77.8%, while the loss was 1.28. This result showed the good potential of CNN models in the automated discrimination of building damage levels. According to the loss at every 100 iterations, the testing loss deviated from the training loss in many models because of the limited number of training data. In the future, more on-site photos will be added to the data sets to improve the performance of the models. Furthermore, the photos which captured parts of a building, e.g. roofs or pillars, will be fed to the models to complete the certification procedure of damage levels.

5. Acknowledgments

This research was supported by "the Tokyo Metropolitan Resilience Project" of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of the Japanese Government, the National Research Institute for Earth Science and Disaster Resilience (NIED), and Niigata University.

6. References

[1] QuiQuake: https://gbank.gsj.jp/QuiQuake/QuakeMap/20160416.en.html

[2] Cabinet Office, Government of Japan: Report of damage situation due to the 2016 Kumamoto earthquake (updated on April 12, 2019). http://www.bousai.go.jp/updates/h280414jishin/pdf/h280414jishin_55.pdf (in Japanese)

[3] Kyushu District Administrative Evaluation Bureau, Ministry of Internal Affairs and Communications: Survey on the issuance of disaster victim certificates during large-scale disasters -Focusing on the 2016 Kumamoto Earthquake-. https://www.soumu.go.jp/main_content/000528758.pdf (in Japanese)

[4] Vetrivel A, Gerke M, Kerle N, Nex F (2017): Vosselman, G. Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, **140**, 45–59.

[5] Duarte D, Nex F, Kerle N, Vosselman, G (2018): Multi-Resolution Feature Fusion for Image Classification of Building Damages with Convolutional Neural Networks. *Remote Sensing*, **10**, 1636.

[6] Nex F, Duarte D, Tonolo FG, Kerle N (2019): Structural Building Damage Detection with Deep Learning: Assessment of a State-of-the-Art CNN in Operational Conditions. *Remote Sensing*, **11**, 2765.

[7] Ueoka Y, Tanaka S (2018): Deep learning interpretation of building damages using aerial and ground photographs. *42nd Conference on Urban Disaster Reduction*, 49–52, Hokkaido, Japan. (in Japanese)

[8] Ishii Y, Matsuoka M, Maki N, Horie K, Tanaka S (2018): Recognition of damaged building using deep learning based on aerial and local photos taken after the 1995 Kobe Earthquake. *Journal of Structural and Construction Engineering*, **83**(751), 1391–1400. (in Japanese)

[9] Cabinet Office, Government of Japan: Guidelines for the damage standards of houses related to earthquakes. http://www.bousai.go.jp/taisaku/pdf/h3003shishin_2.pdf (in Japanese)

[10] Simonyan K, Zisserman A (2015): Very deep convolutional networks for large-scale image recognition. Computer Science. *International Conference on Learning Representations 2015*, San Diego, U. S. A.

[11] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014): Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, **15**(1), 1929–1958.

[12] He K, Zhang X, Ren S, Sun, J (2016): Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, U. S. A.

[13] Caffe: https://caffe.berkeleyvision.org/

[14] Nesterov YE (1983): A method for solving the convex programming problem with convergence rate O(1/k^2). *Dokl. Akad. Nauk SSSR*, **269**, 543–547.

