

TzanckNet: A convolutional neural network to identify cells in the cytology of erosive-vesiculobullous diseases

Mehmet Alican Noyan ^{1, *}, Murat Durdu ², Ali Haydar Eskiocak ²

¹ Ipsumio B.V., High Tech Campus, 5656 AE Eindhoven, The Netherlands

² Department of Dermatology, Başkent University Medical School, Adana Dr. Turgut Noyan Application and Research Center, Adana, Turkey

* Correspondence: alican.noyan@ipsumio.com

Abstract: Tzanck smear test is a low-cost, rapid and reliable tool which can be used for the diagnosis of many erosive-vesiculobullous, tumoral and granulomatous diseases. Currently its use is limited mainly due to lack of experience in interpretation of the smears. We developed a deep learning model, TzanckNet, that can identify cells in Tzanck smear test findings. TzanckNet was trained on a retrospective development dataset of 2260 Tzanck smear images collected between December 2006 – December 2019. The finalized model was evaluated using a prospective validation dataset of 359 Tzanck smear images collected from 15 patients during January 2020. It is designed to recognize six cell types (acantholytic cells, eosinophils, hypha, multinucleated giant cells, normal keratinocytes and tadpole cells). For 359 images and 6 cell types, TzanckNet made 2154 predictions. The accuracy was 94.3 % (95% CI 93.4 to 95.3), the sensitivity was 83.7 % (95% CI 80.3 to 87.0) and the specificity was 97.3 % (95% CI 96.5 to 98.1). The area under the receiver operating characteristic curve was 0.974. Our results show that TzanckNet has the potential to lower the experience barrier needed to use this test, broadening its user base, and hence improving patient well-being.

Keywords: Tzanck Smear test, Deep Learning, Cutaneous Cytology, Vesiculobullous Skin Disease

Introduction

A dermatologist is expected to provide an accurate diagnosis. Accuracy, however, is hardly the only criterion for choosing the right diagnostic test. The cost has always been an important concern and is even more so today with rising healthcare costs in certain countries. Regulations add another dimension to consider. Many tests that are taken for granted in some clinics, might not be available in resource-constrained settings. Patient discomfort is yet another consideration. Therefore, it is hard to find a diagnostic test that checks all the boxes.

Tzanck smear test, a dermatocytological diagnostic tool first described in 1947¹, costs less than a dollar, gives results in less than an hour, causes minimal patient discomfort, does not require anesthesia and is accurate.² It requires simple tools that almost all the clinics in the world can access – a microscope slide, an immersion oil, a scalpel, cytological stains and a microscope. With these advantages, one would expect it to be indispensable for any dermatologist's toolbox.

Surprisingly, the reality is quite the opposite. Although this dermatologist friendly method can be used for the diagnosis of many erosive-vesiculobullous, tumoral and granulomatous diseases, its use is usually limited to herpes virus infections in daily dermatology practice.² In some clinics, this test is not used even for the diagnosis of herpes virus infections; instead patients receive local anesthesia followed by a skin biopsy.³ Furthermore, when folliculitis patients are treated without cytological examination, patients with viral, parasitic and fungal folliculitis may receive unnecessary antibiotic treatments for years.⁴ One of the reasons for this is the lack of experience in using this test. Efforts have been made in the past decade to reintroduce Tzanck smear test⁵⁻¹⁰ which has resulted in revived interest.¹¹⁻¹³ Still, this diagnostic tool is nowhere near its full potential. The main textbooks of dermatopathology do not include dedicated sections for cytology or Tzanck smear.^{14,15}

Artificial intelligence has long been used to tackle challenges in healthcare¹⁶⁻¹⁹, with growing attention in the last years. It has found wide ranging roles from automating menial tasks²⁰ to optimizing medical resources.²¹ The spotlight, however, is on deep learning models showing

expert level performance on certain medical skills – even surpassing humans in some instances.^{22–25} Analyzing medical images constitute an important part of these skills²⁶ because the development of deep learning has been strongly tied to image processing.²⁷ Artificial intelligence models have been used in dermatology^{19,23,28–31}, but, to the best of our knowledge, they have never been applied to cutaneous cytology or evaluating Tzanck smear findings.

The aim of this study was to develop a deep neural network, called TzanckNet, for recognizing cells in Tzanck smear test findings of erosive-vesiculobullous diseases. It was designed to recognize six cell types related to diseases such as herpetic infections, pemphigus, and spongiotic dermatitis (Figure 1). The model was developed using a retrospective dataset and validated clinically using a prospective dataset.

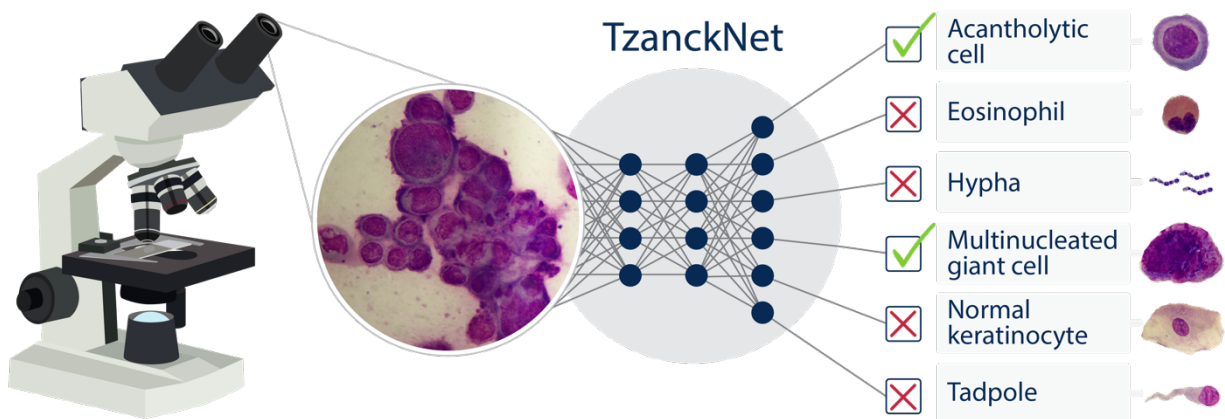


Figure 1. A schematic showing how the proposed TzanckNet works. TzanckNet accepts a Tzanck smear image as an input and outputs which cell types among six cell types are present and absent in the image. As an example, a Tzanck smear image from a patient with herpetic infection goes into the TzanckNet and the network predicts that there are acantholytic and multinucleated giant cells in the image, the remaining four cell types does not exist in the image. On the right-hand side of the figure, an example image from each cell type is presented.

Methods

Datasets. For developing the TzanckNet, we used the cytology archive of Department of Dermatology, Başkent University Medical School, Adana Dr. Turgut Noyan Application and Research Center. This study was approved by Başkent University Institutional Review Board (Project no: KA19/401). There were 2260 Tzanck smear images from erosive-vesiculobullous diseases (pemphigus, herpetic infection, impetigo, Hailey-Hailey disease, contact dermatitis, vesiculobullous dermatophytic infection) collected between December 2006 – December 2019. This dataset is called the *development dataset*. It was used to train and fine tune the model. Afterwards, the finalized model was evaluated using a separate dataset that was collected during January 2020 in the same clinic. It contained 359 images from patients that were not in the development dataset. This is called the *validation dataset*. The images in this dataset were obtained from 15 patients (9 females, 6 males) with ages ranging from 2 to 68 years (34 years on average).

Dataset preparation. Cytological specimens were stained with May-Grünwald-Giemsa (Bio-Optica), examined microscopically and photographed with a digital camera. Cytological photos taken at x1000 magnification. Two dermatologists (M.D. and A.H.E.) prepared the dataset with two different microscopes and with two different digital cameras. There were no missing data in the datasets. Images with artifacts (overlapping cells, scratches etc.) were not removed from the dataset. Only images that were blurry at a level that would make them useless in a clinical setting were discarded and the samples were imaged again. Images were labeled, de-identified and given to the data scientist (M.A.N.).

Reference standard. Images were labeled by two dermatologists (M.D. and A.H.E.). M.D. has 13 years and A.H.E. has 2 years of previous cutaneous cytology experience. Labels were decided by adjudication³² i.e. disagreements were resolved through discussion. The image was labeled positively for a cell only if that cell was entirely present in the image.

Machine learning model development. A deep learning library, fastai³³, was used to develop the TzanckNet. It is a convolutional neural network (ResNet-50), pretrained on ImageNet. Then, the final classification layer was replaced with six output nodes and retrained. Six output nodes in the final layer correspond to the six cell types we chose our model to recognize: acantholytic cells, eosinophils, hypha, multinucleated giant cells, normal keratinocytes and tadpole cells (Figure 1). An image can contain any combination of these cells; therefore, TzanckNet was designed such that it can output any number/combination of cells for one input image. Each output node, outputs a value between 0 and 1, corresponding to the probability of that cell existing in the input image. Depending on the discrimination threshold this probability is converted to 1 or 0 indicating whether the cell type exists in the image or not. For this study, the discrimination threshold was kept constant as 0.5, except for plotting the receiver operating characteristic curve. The development dataset was used at this stage, for training and tuning the model.

Statistical analysis. Discrimination and calibration performance were used to assess the TzanckNet. The following metrics were used for discrimination: accuracy, sensitivity, specificity, positive predictive value, negative predictive value and F_1 score with a discrimination threshold of 0.5. Additionally, receiver operating characteristic curve (ROC curve) was plotted to evaluate the model at different discrimination threshold values. Area under this curve (AUC) was calculated and reported as well. Calibration curve was used to evaluate how well predicted probabilities approximated the actual event probabilities. The validation dataset was used at this stage for evaluating the model. Finally, TzanckNet predictions on selected images were demonstrated.

Results

Using a convolution neural network, we created a machine learning model called TzanckNet for recognizing six cell types in a Tzanck Smear image. An image can contain multiple cell types, for this reason we designed the model such that it can output more than one cell type. Given one input image, the model outputs six probabilities, corresponding to the predicted probabilities of the existence of six cell types. We fixed a discrimination threshold of 0.5 such that if the predicted

probability for any cell type was above this threshold, the model predicted that this cell type existed in the image. These predictions were compared to the reference standard for evaluating the model performance.

The TzanckNet, trained and tuned on the development set, was prospectively validated using 359 images collected in a real-world setting. For each image it made six predictions hence it made 2154 predictions in total. The overall accuracy of the TzanckNet on this dataset was 94.3 % (95% CI 93.4 to 95.3), the sensitivity was 83.7 % (95% CI 80.3 to 87.0) and the specificity was 97.3 % (95% CI 96.5 to 98.1). Results of the other discrimination metrics and performance for each cell are given in Table 1.

Table 1. Discrimination metrics of the TzanckNet on the validation dataset that contains 359 Tzanck smear findings.

	Accuracy	Sensitivity	Specificity	PPV ^a	NPV ^b	AUC ^c	F ₁ score
Acantholytic cell	315/359 (87.7 %)	118/139 (84.9 %)	197/220 (89.5 %)	118/141 (83.7 %)	197/218 (90.4 %)	0.954	84.3 %
Eosinophil	324/359 (90.3 %)	24/59 (40.7 %)	300/300 (100 %)	24/24 (100 %)	300/335 (89.6 %)	0.918	57.8 %
Hypha	352/359 (98.1 %)	39/39 (100 %)	313/320 (97.8 %)	39/46 (84.8 %)	313/313 (100.0 %)	0.999	91.8 %
Multinucleated giant cell	353/359 (98.3 %)	109/114 (95.6 %)	244/245 (99.6 %)	109/110 (99.1 %)	244/249 (98.0 %)	0.998	97.3 %
Normal keratinocyte	359/359 (100 %)	47/47 (100 %)	312/312 (100 %)	47/47 (100 %)	312/312 (100 %)	1	100 %
Tadpole cell	329/359 (91.6 %)	52/67 (77.6 %)	277/292 (94.9 %)	52/67 (77.6 %)	277/292 (94.9 %)	0.959	77.6 %
Overall	2032/2154 (94.3 %)	389/465 (83.7 %)	1643/1689 (97.3 %)	389/435 (89.4 %)	1643/1719 (95.6 %)	0.974	86.4 %

^a PPV: Positive predictive value ^b NPV: Negative predictive value ^c AUC: Area under the receiver operating characteristic curve.

In order to visualize model performance across different discrimination thresholds, overall receiver operating characteristic curve of the network is given in Figure 2a. The area under this curve (AUC) was 0.974. For assessing the performance of the predicted probabilities, calibration curve is given in Figure 2b.

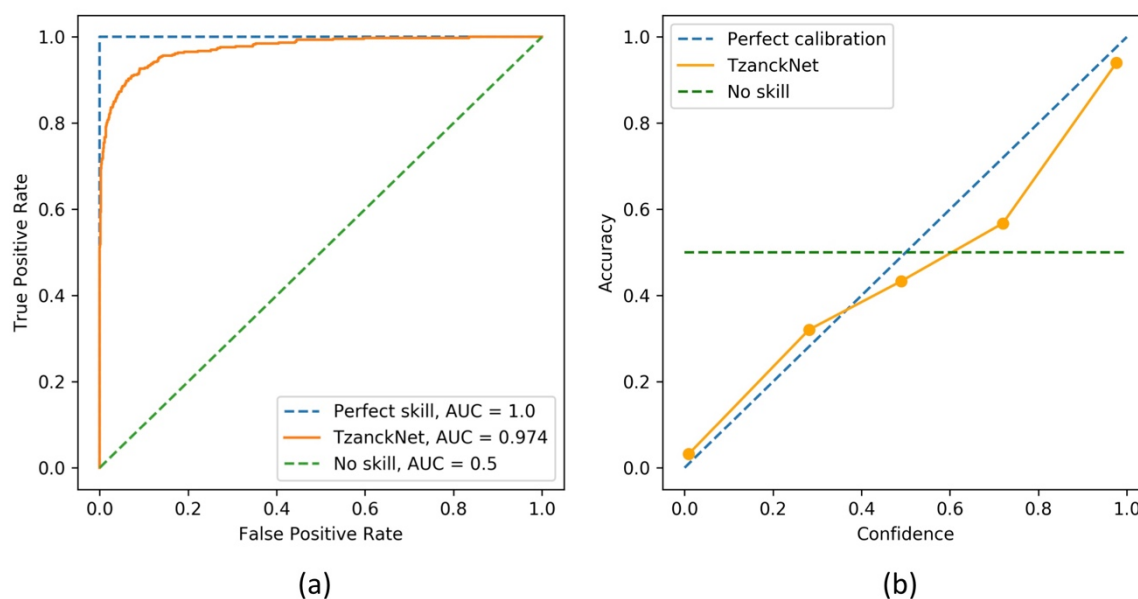


Figure 2. (a) Receiver operating characteristic (ROC) and (b) calibration curves of the TzanckNet on the validation set.

TzanckNet performance on selected Tzanck smear findings are given in Figure 3. Figure 3a contains a multinucleated giant cell (arrow). At first sight during labeling, it seemed like the cell had one circular nucleus and was hence labeled as an acantholytic cell. Upon further examination, it was realized that the cell actually had three nuclei instead of one and the label was changed to a multinucleated giant cell. The TzanckNet recognized that there was as a multinucleated giant cell, but it also gave a 44.6 % probability for the existence of an acantholytic cell. Two eosinophils can be seen in Figure 3b (arrows), however, TzanckNet didn't recognize them. In Figure 3c, only some part of an acantholytic cell (arrow) is visible inside the image therefore the image was labelled as having no acantholytic cell in the image. TzanckNet predicted that there was an acantholytic cell, but it was not confident (51.1 % predicted probability). Figure 3d contains an

atypical acantholytic cell (arrow). It is atypical because the perinuclear halo is not evident and in some parts cytoplasm is not discernible. Therefore, the image was labeled as not having an acantholytic cell. The TzanckNet, however, detected that there is an acantholytic cell in this image, and it was 86.3 % confident.

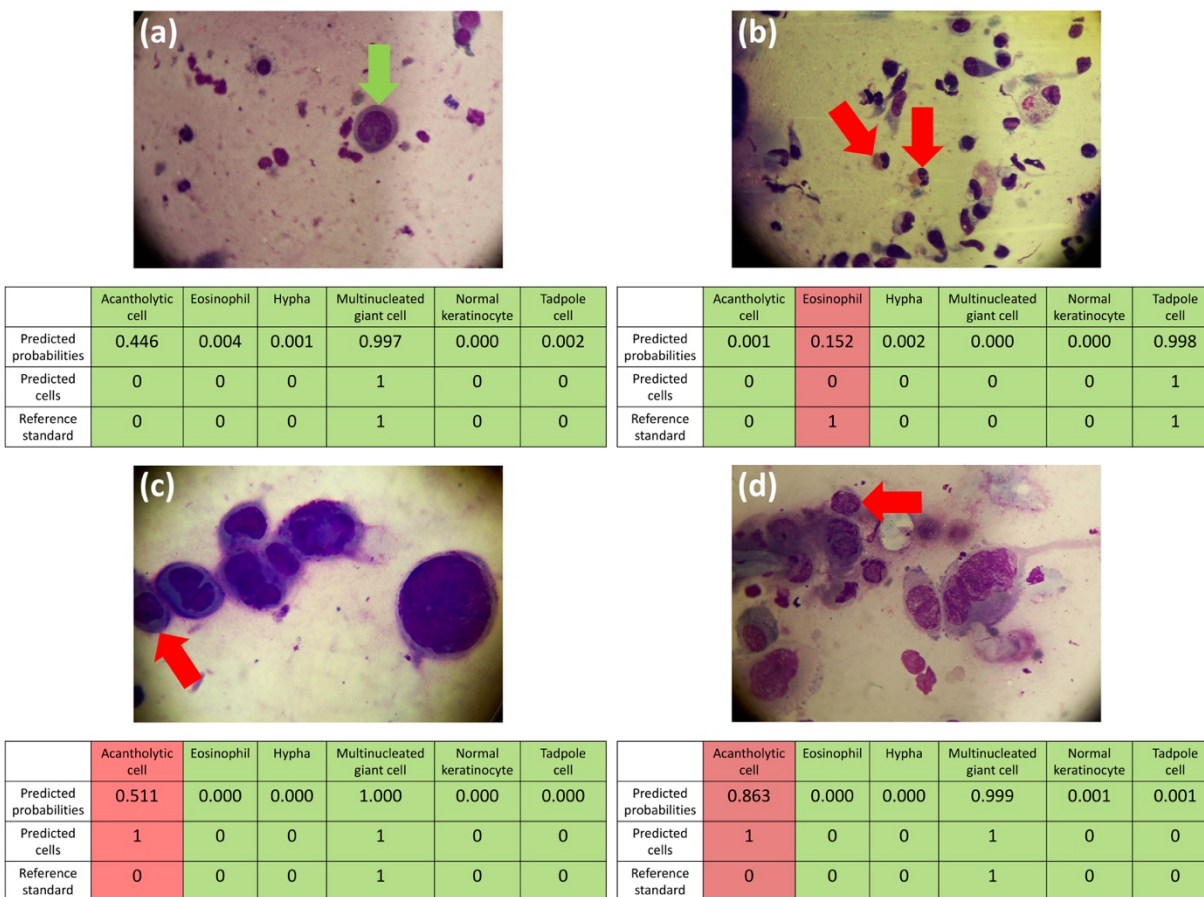


Figure 3. TzanckNet predictions and the corresponding reference standards for four selected images. For each cell type and image, TzanckNet predicts the probability of that cell type being present in the image. Probabilities are then converted to decisions of absence (0) or presence (1) of that cell, using a discrimination threshold of 0.5. Decisions matching with the reference standards are marked with green, and with red otherwise. The red arrows indicate the cells that are related to the false predictions. The green arrow indicates a multinucleated giant cell.

Discussion

Tzanck smear test is an inexpensive test that can provide rapid diagnosis of many dermatological diseases. Interest in Tzanck smear test has increased in recent years. Nevertheless, in most clinics, this simple test is either not performed or is only used to diagnose a few diseases, mainly due to lack of experience. In a dermatology clinic where 75 patients for whom no diagnosis could be established via face-to-face clinical examinations and cytological evaluations, telemedicine was used to overcome this issue.³⁴ Although this is a reasonable solution, the necessary expertise is still scarce, scheduling is necessary, and one can only examine a limited number of images during a session. In this work we demonstrated that TzanckNet can identify six cell types in Tzanck smear findings. It can analyze hundreds of images in a minute with high accuracy.

The overall performance of the TzanckNet shows that deep learning has the potential to aid dermatologists for analyzing Tzanck smear findings. It was able to recognize cells with high accuracy. Moreover, the calibration performance shows that the predicted probabilities can also help with the interpretation of images. For example, if a cell type exists with 99 % estimated probability the network is almost certain that this cell is present in the image whereas if the estimated probability is 70 % the network tends to predict that this cell is present but not confidently. If the network is not confident about its predictions, additional images can be taken.

Examining individual predictions can provide a better understanding of the model's strengths and limitations hence indicating ways to improve it furthermore. For example, TzanckNet was able to recognize correctly a multinucleated cell that looks similar to an acantholytic cell (Figure 3a). Looking at Table 1, low sensitivity on eosinophils draws attention. Specificity being 100 % indicates that all of the errors for eosinophils were false negatives, there were no false positives. Individual examination of the images that resulted in false negatives revealed that they had less than five eosinophils on average whereas the examination of true positives revealed that they had more than ten eosinophils on average. One of the images that resulted in a false negative eosinophil can be seen in Figure 3b. Indeed, it contains only two eosinophils (arrows). This suggests that for TzanckNet to recognize eosinophils, they should be abundant in the image.

Figure 3c and 3d indicates model performance may not be ideal for cells that are partially in the image or atypical cells.

The most common use of Tzanck smear test is the diagnosis of herpetic infection.^{7,35} Herpetic infections may be primary or develop on other dermatoses. Early diagnosis of herpetic infections is very important in terms of early treatment and prevention of the spread of the disease. When Tzanck smear is not performed, patients receive unnecessary treatment or more complex diagnostic methods such as polymerase chain reaction or histopathological examinations are requested.² Tzanck smear test is cheaper and faster than both of these methods.⁴ The characteristic finding of herpetic infection in the Tzanck smear test is acantholytic cells with multinuclear giant cells.² Another disease that is related to these cells is pemphigus. It is a rare and severe autoimmune disease that can cause mortality. Tzanck smear test is not only important in early diagnosis of the disease but also in the diagnosis of recurrent lesions. A definitive diagnosis of pemphigus requires histopathological examination and immunofluorescence test. However, it is not possible to take a biopsy each time for recurrent lesions. Herpetic infection and candidiasis should be excluded ruled out especially in newly developed lesions resistant to corticosteroid treatment in oral mucosa. If this test is not performed, patients will receive unnecessary steroid treatment, which may lead to sepsis.³⁶ In pemphigus patients, unlike herpetic infections, acantholytic cells without multinuclear giant cells are observed in cytological examination.^{2,9} In this study, TzanckNet showed that it was able to recognize acantholytic and multinuclear giant cells demonstrating its diagnostic potential for suspected herpetic infection and pemphigus.

A few tadpole cells are observed in all vesiculobullous diseases. However, numerous tadpole cells are characteristic cytologic findings of spongiotic dermatitis³⁷ and eczema. Tzanck smear is used for two purposes in eczema patients. The first is to distinguish eczema from other acantholytic diseases and the second is to detect bacterial, fungal and viral infections. Tzanck smear test performed for both purposes prevent unnecessary treatment. If the test is not done, infections can spread easily with steroid treatments used for eczema. Accuracy of the model on tadpole

cells was 91.6 %, however it should be noted that the specificity was high, but sensitivity was low suggesting that the model tends to give false negatives.

Cutaneous eosinophilic infiltration may develop due to various causes such as infectious, inflammatory and neoplastic diseases. Histopathology is often used in the differentiation of these diseases. However, cytological examination has been used in only a few studies focused on the utility of cytology in some eosinophilic diseases. Besides detecting eosinophils, cytology can also show the various infectious etiologic agents and distinguish some inflammatory diseases. If eosinophils are present, bacterial and fungal structures should be investigated. The performance of TzanckNet on eosinophils and hypha suggests that the network can also be utilized as a decision support tool for diagnosing aforementioned diseases.

The overall performance of the TzanckNet demonstrates its strong diagnostic potential. We would like to emphasize that we propose this model as a clinical decision support system to improve physician accuracy and efficiency in the clinical workflow, not as a replacement or a tool that can directly output a diagnosis. It has been shown that human-machine combination works better than either alone.³⁸⁻⁴⁰ Additionally, TzanckNet can be used to train dermatologists on Tzanck smear test.

Limitations. Like all methods/models, explicitly reporting known limitations of TzanckNet is crucial for clinicians and researchers who want to use and improve it. The model was trained using Tzanck smear findings of patients from Turkey, stained with May-Grünwald-Giemsa and taken with x1000 magnification. The performance of the model on other populations, sample preparation procedures and imaging systems remains unknown at this point. Model performance on cells that are atypical or that overlap or appear partially in the image may not be ideal. Finally, yet importantly, the model can recognize which cell types exist in an image, but it cannot localize the cells. Localizing the cells has many advantages such as being able to count the cells and being able to say precisely which cells in the image are recognized and what their types are. Two major improvements are needed to realize the full potential of the TzanckNet: the first is to retrain and

validate the model with other populations, stains, and magnifications for better generalization, and the second is to add object detection to the model for localizing the cells.

Nevertheless, this study showed that TzanckNet can analyze the cytological findings of erosive-vesiculobullous diseases accurately. The model can be extended for analyzing granulomatous and tumoral diseases.

Conclusions

Tzanck smear test is a valuable but underappreciated diagnostic tool. This work introduced TzanckNet, a machine learning model that can analyze Tzanck smear findings with high accuracy. It can be used as a clinical decision support system as well as a training tool for new physicians. It has the potential to spread the use of Tzanck smear test, decrease the number of biopsies, prevent unnecessarily long antibiotic treatments, help early diagnoses for fatal diseases, decrease costs, and thus improve patient well-being.

Author contributions. M.A.N. and M.D. had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. M.A.N. and M.D. proposed the study and its design. M.D. and A.H.E. prepared, imaged, and labeled the Tzanck smears. M.A.N. developed the machine learning model and wrote the manuscript. All authors reviewed, revised, and approved the final manuscript.

Funding. None

Conflict of Interest Disclosures. None

References

1. Tzanck A. Le cyto-diagnostic immédiat en dermatologie. *Press Medicale*. 1947;55(25):286.
2. Durdu M. *Cutaneous Cytology and Tzanck Smear Test*. 1st ed. Springer International Publishing; 2019. doi:10.1007/978-3-030-10722-2
3. Horn TD. Commentary: Heading the wrong way. The disappearing Tzanck smear. *J Am Acad Dermatol*. 2008;59(6):965-966. doi:10.1016/j.jaad.2008.08.025
4. Durdu M, Ilkit M. First step in the differential diagnosis of folliculitis: Cytology. *Crit Rev Microbiol*. 2013;39(1):9-25. doi:10.3109/1040841X.2012.682051
5. Ruocco V, Ruocco E. Tzanck smear, an old test for the new millennium: When and how. *Int J Dermatol*. 1999;38(11):830-834. doi:10.1046/j.1365-4362.1999.00815.x
6. Gupta LK, Singhi MK. Tzanck smear: A useful diagnostic tool. *Indian J Dermatol Venereol Leprol*. 2005;71(4):295-299. doi:10.4103/0378-6323.16632
7. Durdu M, Baba M, Seçkin D. The value of Tzanck smear test in diagnosis of erosive, vesicular, bullous, and pustular skin lesions. *J Am Acad Dermatol*. 2008;59(6):958-964. doi:10.1016/j.jaad.2008.07.059
8. Kelly B, Shimoni T. Reintroducing the Tzanck smear. *Am J Clin Dermatol*. 2009;10(3):141-152. doi:10.2165/00128071-200910030-00001
9. Durdu M, Baba M, Seçkin D. More experiences with the Tzanck smear test: Cytologic findings in cutaneous granulomatous disorders. *J Am Acad Dermatol*. 2009;61(3):441-450. doi:10.1016/j.jaad.2009.02.050
10. Ruocco E, Brunetti G, Del Vecchio M, Ruocco V. The practical use of cytology for diagnosis in dermatology. *J Eur Acad Dermatology Venereol*. 2011;25(2):125-129. doi:10.1111/j.1468-3083.2010.03740.x
11. Yaeen A, Ahmad Q, Farhana A, Shah P, Hassan I. Diagnostic value of Tzanck smear in various erosive, vesicular, and bullous skin lesions. *Indian Dermatol Online J*. 2015;6(6):381. doi:10.4103/2229-5178.169729
12. Zhou T, Fang S, Li C, Hua H. Comparative study of indirect immunofluorescence, enzyme-linked immunosorbent assay, and the Tzanck smear test for the diagnosis of pemphigus. *J*

- Oral Pathol Med.* 2016;45(10):786-790. doi:10.1111/jop.12439
13. Panwar H, Joshi D, Goel G, Asati D, Majumdar K, Kapoor N. Diagnostic utility and pitfalls of Tzanck smear cytology in diagnosis of various cutaneous lesions. *J Cytol.* 2017;34(4):179-182. doi:10.4103/JOC.JOC_88_16
 14. Elder D. *Lever's Histopathology of the Skin.* 11th ed. Wolters Kluwer; 2014. doi:9781451190373
 15. Barnhill R, Crowson N, Magro C, Piepkorn M, Kutzner H, Desman G. *Dermatopathology.* 4th ed. McGraw-Hill Education; 2019.
 16. Bartels PH, Weber JE, Duckstein L. Machine learning in quantitative histopathology. *Anal Quant Cytol Histol.* 1988;10(4):299-306.
 17. Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med.* 1992;21(4):384-390. doi:10.1016/S0196-0644(05)82656-3
 18. Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computer-derived nuclear features distinguish malignant from benign breast cytology. *Hum Pathol.* 1995;26(7):792-796. doi:10.1016/0046-8177(95)90229-5
 19. Dreiseitl S, Ohno-Machado L, Kittler H, Vinterbo S, Billhardt H, Binder M. A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *J Biomed Inform.* 2001;34(1):28-36. doi:10.1006/jbin.2001.1004
 20. Lin SY, Shanafelt TD, Asch SM. Reimagining Clinical Documentation With Artificial Intelligence. *Mayo Clin Proc.* 2018;93(5):563-565. doi:10.1016/j.mayocp.2018.02.016
 21. Horng S, Sontag DA, Halpern Y, Jernite Y, Shapiro NI, Nathanson LA. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. Groza T, ed. *PLoS One.* 2017;12(4):e0174708. doi:10.1371/journal.pone.0174708
 22. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA - J Am Med Assoc.* 2016;316(22):2402-2410. doi:10.1001/jama.2016.17216
 23. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115-118. doi:10.1038/nature21056

24. Chilamkurthy S, Ghosh R, Tanamala S, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet*. 2018;392(10162):2388-2396. doi:10.1016/S0140-6736(18)31645-3
25. Tomašev N, Glorot X, Rae JW, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature*. 2019;572(7767):116-119. doi:10.1038/s41586-019-1390-1
26. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88. doi:10.1016/j.media.2017.07.005
27. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. doi:10.1038/nature14539
28. Fujisawa Y, Otomo Y, Ogata Y, et al. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *Br J Dermatol*. 2019;180(2):373-381. doi:10.1111/bjd.16924
29. Tschandl P, Rosendahl C, Akay BN, et al. Expert-Level Diagnosis of Nonpigmented Skin Cancer by Combined Convolutional Neural Networks. *JAMA Dermatology*. 2019;155(1):58-65. doi:10.1001/jamadermatol.2018.4378
30. Brinker TJ, Hekler A, Enk AH, et al. A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. *Eur J Cancer*. 2019;111:148-154. doi:10.1016/j.ejca.2019.02.005
31. Han SS, Moon IJ, Lim W, et al. Keratinocytic Skin Cancer Detection on the Face Using Region-Based Convolutional Neural Network. *JAMA Dermatology*. 2020;156(1):29-37. doi:10.1001/jamadermatol.2019.3807
32. Krause J, Gulshan V, Rahimy E, et al. Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy. *Ophthalmology*. 2018;125(8):1264-1272. doi:10.1016/j.ophtha.2018.01.034
33. Howard J, Gugger S. Fastai: A Layered API for Deep Learning. *Information*. 2020;11(2):108. doi:10.3390/info11020108
34. Durdu M, Harman M. Diagnostic value of telecytology in tertiary teledermatological consultation: a retrospective analysis of 75 cases. *Int J Dermatol*. 2016;55(7):e392-e398.

doi:10.1111/ijd.13219

35. Solomon AR, Rasmussen JE, Varani J, Pierson CL. The Tzanck Smear in the Diagnosis of Cutaneous Herpes Simplex. *JAMA J Am Med Assoc.* 1984;251(5):633-635. doi:10.1001/jama.1984.03340290047019
36. Kalajian AH, Callen JP. Atypical herpes simplex infection masquerading as recalcitrant pemphigus vulgaris. *Australas J Dermatol.* 2007;48(4):242-247. doi:10.1111/j.1440-0960.2007.00406.x
37. Pariser RJ. Diagnosis of spongiotic vesicular dermatitis by Tzanck smear: The “tadpole cell.” *J Am Acad Dermatol.* 1983;8(4):519-522. doi:10.1016/S0190-9622(83)70058-7
38. Steiner DF, Macdonald R, Liu Y, et al. Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am J Surg Pathol.* 2018;42(12):1636-1646. doi:10.1097/PAS.0000000000001151
39. Sayres R, Taly A, Rahimy E, et al. Using a Deep Learning Algorithm and Integrated Gradients Explanation to Assist Grading for Diabetic Retinopathy. *Ophthalmology.* 2019;126(4):552-564. doi:10.1016/j.ophtha.2018.11.016
40. Patel BN, Rosenberg L, Willcox G, et al. Human-machine partnership with artificial intelligence for chest radiograph diagnosis. *npj Digit Med.* 2019;2(1):111. doi:10.1038/s41746-019-0189-7