УДК 378:004
DOI: 10.25686/978-5-8158-2474-4-2025-835-843

Изучение нескольких ResNet-подобных моделей для классификации атипичных митотических фигур

Александр Трыкин[1,2], Евгений Васильев[1], Вадим Турлапов[1]
[1]Исследовательский центр искусственного интеллекта, Институт информационных технологий, математики и механики, Нижний Новгород, Россия
[2]Лаборатория системной медицины старения, Научно-исследовательский институт биологии старения, Нижний Новгород, Россия

Аннотация. Атипичные изменения в клеточном делении могут указывать на опасные патологические состояния, связанные с высокоагрессивными формами злокачественных новообразований. Идентификация митотических фигур — субъективная и трудоёмкая задача в гистопатологии. В данной работе представлены ResNet-подобные модели для решения задачи классификации нормальных/атипичных митотических фигур. Были взяты два общедоступных набора данных AMi-Br и MIDOG25, посвящённых этой относительно новой проблеме, и объединены, чтобы максимально использовать имеющиеся данные и охватить различные условия. Рассмотрено использование нескольких функций потерь, таких как Focal Loss (FL), Recall Loss (RL) и Focal Recall Loss (FRL). Наилучшие значения сбалансированной точности 84.93 % и AUROC 92.12 % продемонстрировала модель ResNetV2-34D с Recall Loss, F1 76,81 % показала ResNeSt-26D также с Recall Loss.
Ключевые слова: атипичный митоз, гистопатология, глубокое обучение, несбалансированное обучение

Exploring several ResNet-like models for atypical mitotic figure classification

Alexander Trykin[1,2], Evgenii Vasiliev[1] and Vadim Turlapov[1]
[1]Artifcial Intelligence Research Center, Institute of Information Technologies, Mathematics and Mechanics, Lobachevsky State University, Nizhny Novgorod, Russia
[2]Laboratory of Systems Medicine of Ageing, Research Institute for Biology of Aging, Lobachevsky State University, Nizhny Novgorod, Russia

Abstract. Atypical changes in cell division may indicate dangerous pathological conditions associated with highly aggressive forms of malignant neoplasms. Identification of mitotic figures is a subjective and labour-intensive task in histopathology. In this paper, we present ResNet-like models for solving the problem of normal/atypical mitotic figure classification. We took two publicly available datasets devoted to this relatively new problem: AMi-Br and MIDOG25, and combined them to use as much available data as possible and cover different conditions. We have looked at the use of several loss functions, such as Focal Loss (FL), Recall Loss (RL), and Focal Recall Loss (FRL). We were able to achieve the best values of balanced accuracy of 84.93% and AUROC of 92.12% was shown by the ResNetV2-34D model with Recall Loss, and F1 of 76.81% was ResNeSt-26D also with Recall Loss.
Keywords: atypical mitosis, histopathology, deep learning, imbalanced learning

Introduction

Cell division is a fundamental biological process, ensuring the growth, development, and maintenance of multicellular organisms. A critical aspect of this process is mitosis, where genetic material is equally distributed between daughter cells. Under normal conditions, mitotic figures exhibit a standardized morphology, reflecting the highly regulated orchestration of chromosomal and cellular events. However, in certain pathological states, such as cancer, mitotic figures can deviate from their typical appearance, giving rise to what are commonly referred to as atypical mitotic figures (AMF's). These anomalies often signal chromosomal instability, a hallmark of tumorigenesis, and are associated with highly aggressive malignancies, poor prognoses, and resistance to therapeutic interventions [1, 2, 3].

While traditional histopathology has played a significant role in identifying AMF's, the process is often limited by subjectivity and requires significant expertise. Recent advancements in imaging, artificial intelligence (AI), and computational pathology have opened new pathways for automated, precise, and reproducible classification of atypical mitotic figures [4, 5, 6].

The recent Mitosis Domain Generalization Challenge 2025 (MIDOG25) [7] provides a standardized benchmark for creating accurate and reliable solutions, calling on researchers to contribute to the development of AI cancer assessment tools to assist pathologists. The authors of the benchmark note several key issues: (1)

severe class imbalance, (2) high morphological variability within each class, (3) little difference between normal and atypical figures, and (4) a wide range of different tumours, types, staining methods, and scan areas.

Although the challenge is relatively new, work on this topic is just beginning to emerge. The authors of the original Ami-Br dataset presented results for their dataset [8]. They tried to deal with imbalance using various techniques such as weighted cross-entropy, focal loss, and weighted sampling. They tried two deep learning models: DenseNet-121 and EfficientNet V2-S. They report achieving an average balanced accuracy of around 80.6%. In [9] authors presented two new datasets: AtNorM-Br, a mitotic figure dataset from the TCGA breast cancer cohort, and AtNorM-MD, a multi-domain mitotic figure dataset from aч subset of the MIDOG++ training set. They tried approaches including end-to-end trained deep learning models, foundation models with linear probing, and foundation models fine-tuned with low-rank adaptation (LoRA). They report achieving 81.35%, 77.88%, and 77.23% average balance accuracy on the in-domain AMi-Br and the out-of-domain AtNorm-Br and AtNorM-MD datasets, respectively

In this paper, we consider several ResNet-like models for the atypical mitotic figure classification task. We use two public datasets, AMi-Br and MIDOG25, to make the most of the available data and thus cover different tumours, tissues, scanners, i.e., very different conditions. We use on-the-fly data augmentation and weighted random sampling to combat class imbalance. Moreover, we also test several loss functions, such as Focal Loss (FL), Recall Loss (RL), and Focal Recall Loss (FRL).

Materials

1. Used Datasets

We utilized two publicly available datasets, AMi-Br [8] and MIDOG25 [10], which were specifically designed to support the development and evaluation of mitotic figure classification methods. These datasets consist of annotated histopathological images of tissue samples from patients with tumours originating from a variety of organ sites. More detailed information about each dataset is given in Tab. 1 and below:

Table 1. **Datasets description**

Property	AMi-Br	MIDOG25
Image Size	128 x 128	128 x 128
Total MFs	3720	11939
Atypical	832	1748
Normal	2888	10191
AMF Rate (%)	22.4%	14.6%
Annotation Type	3-expert vote	3-expert vote
Expert Agreement	78.2%	70%
Source Datasets	TUPAC16, MIDOG21	MIDOG++ (MIDOG21, MIDOG22)
Species	Human	Human + Canine

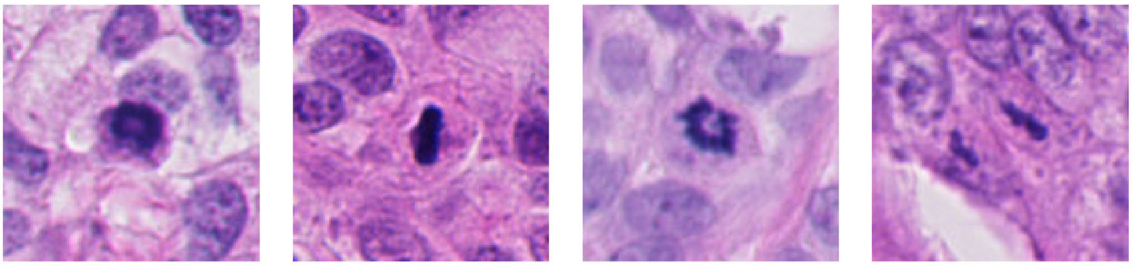
- AMi-Br: This dataset consists of two previously largest and most diverse human breast cancer challenges, namely the TUPAC [11] using an improved alternative version of annotations and the MIDOG 2021 [12]. The samples were processed in several laboratories: UMC Utrecht, Symbiant Pathology Expert Center (Alkmaar & Zaandam), and using several scanners: Hamamatsu NanoZoomer XR (C1200022), Hamamatsu NanoZoomer S360, Aperio Scanscope CS2, Aperio ScanScope XT, Leica Aperio GT 450, Leica SCN400. All histological sections were stained with standard H&E stain and scanned using a 40× objective, giving a scanning resolution of 0.23, 0.25 or 0.26 $\mu m/px$ depending on the scanner. Three expert pathologists were recruited to create the annotations, independently of each other.
- MIDOG25: The MIDOG25 dataset is part of the Mitosis Domain Generalization Challenge 2025 (task 2), created to benchmark AI models for mitotic figure classification under varying imaging conditions and biological domains. Cases of three tumour types were obtained from human patients (breast carcinoma, pancreatic and gastrointestinal neuroendocrine tumours, and cutaneous melanoma) and four tumour types were obtained from dog patients (pulmonary carcinoma, lymphosarcoma mostly in lymph nodes, cutaneous mast cell tumours, and (sub)cutaneous soft tissue sarcoma). The samples were processed in several laboratories: UMC Utrecht, VMU Vienna, FU Berlin, AMC New York, and using several scanners: Hamamatsu XR

(C12000-22), Hamamatsu S360 (0.5 N/A), Leica ScanScope CS2, 3DHistech Pannoramic Scan II, Aperio ScanScope CS2. All histological sections were stained with standard H&E stain and scanned using a 40× objective, giving a scanning resolution of 0.23 or 0.25 $\mu m/px$ depending on the scanner. Three expert pathologists were recruited to create the annotations, who had extensive experience in annotating mitotic figures.

Importantly, the datasets were curated to represent diverse imaging conditions, emulating real-world variability and enabling the development of robust, domain-generalizable classification approaches.

2. Short Data Overview

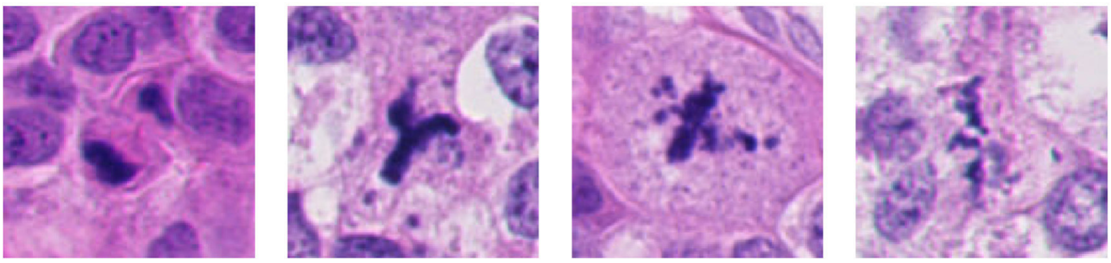
Mitosis is the division of one cell nucleus, resulting in the formation of two daughter nuclei with identical genetic information [1]. Mitotic figures (MF) are structures that can be identified using light microscopy. MFs are defined as nuclear aggregates containing short rods or spikes of chromosomes. Depending on the phase of mitosis, the morphology of MFs varies. The phases of normal mitosis include prophase, prometaphase, metaphase, anaphase, and telophase. It is important to note that mitosis is a continuous process, and cells observed in histological preparations may pause during the transition between phases. The duration of each phase may vary depending on the animal species and cell type. Examples of normal mitotic figures are shown in fig. 1.



(a) prometaphase (b) metaphase (c) ring-shape metaphase (d) anaphase-telophase

Figure 1. Examples of normal mitotic figures (NMF) taken from the presented datasets: (a)–(d) represent different phases of mitosis in which mitotic figures are normal

The process of cell division can be prone to errors in chromosome segregation. When these errors occur in neoplastic (cancerous) cells, they result in genetic abnormalities that are morphologically detectable as AMF. These abnormalities can be broadly classified into two groups: (1) issues related to mitotic or polar asymmetry, and (2) improper segregation of chromosomes. More intricate genetic aberrations can develop from the interplay of these fundamental categories. Examples of atypical mitotic figures are shown in fig. 2.



(a) bipolar asymmetric (b) multipolar (c) segregation (d) other

Figure 2. Examples of atypical mitotic figures (AMF) taken from the presented datasets: (a)–(d) represent various anomalies of mitosis in which mitotic figures are atypical

Methods

1. Evaluation Metrics

Since the problem is unbalanced in terms of data (only about 20 % are atypical figures), it is important to use appropriate metrics for proper evaluation. We chose balanced accuracy as our main metric. We also consider the F1 score and area under ROC curve for more information on testing.

Balanced Accuracy (BA): is a metric designed to address imbalanced datasets by accounting for class-specific performance. It calculates the average of recall values across all classes, ensuring equal importance is given to each class regardless of its prevalence. It can be evaluated in following way:

$$BA = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{FN}{FN + FP}\right). \quad (1)$$

F1 Score (F1): is a harmonic mean of precision and recall, offering a single metric that reflects a model's ability to balance false positives and false negatives. It calculated as:

$$BA = \frac{2TP}{2TP + FP + FN}. \quad (2)$$

Area Under ROC Curve (AUROC): evaluates a classification models ability to distinguish between classes at various decision thresholds. It plots the true positive rate (sensitivity) against the false positive rate and computes the area under this curve. AUC-ROC values range from 0 to 1, where a value closer to 1 signifies excellent discriminative ability. This metric is particularly effective when comparing model performance across different threshold settings or when dealing with imbalanced datasets, as it evaluates performance independent of a specific decision boundary

2. Convolutional Neural Networks

We decided to focus on several ResNet-like models, which are deep convolutional neural networks. ResNet still maintains several advantages and may well be the optimal choice, particularly in certain scenarios. Architectures like ResNet are computationally efficient compared to transformers such as Vision Transformers (ViT) and MLP-Mixers, especially for small and medium datasets or in resource-constrained situations [13]. Transformers are also often more sensitive to hyperparameter tuning. We would also like to focus on simpler and more well-known models first.

A detailed description of each model's architecture is provided below.

2.1. ResNet & ResNetV2

It is known that if you simply increase the number of layers in the network, it begins to learn worse. This result is counterintuitive, as the intuitive assumption is that increased network depth should linearly enhance representational capacity. The fact is that one of the big problems of deep networks is the problem of vanishing gradients [14].

The organization of the network in the form of residual blocks, invented by a team from the Asian division of Microsoft Research, made it possible to significantly increase the depth of trained networks in 2015 [15, 16]. Their idea is to add a shortcut (skip connection) that skips one or more layers.

At the same time, short connections have a positive effect on the landscape of the error function [17, 18]. This does not make the function convex, which would significantly simplify the learning process, but allows the surface to become smoother, which means that gradient descent and its modifications will work better.

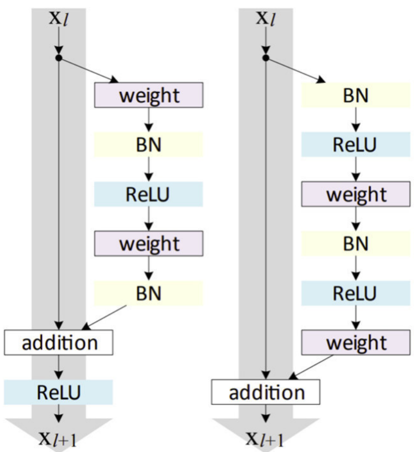


Figure 3. Differences between ResNet (*left*) and ResNetV2 (*right*)

While the original ResNet paper did not draw direct inspiration from biological systems, subsequent research has explored parallels between ResNet's architecture and certain biological learning mechanisms. Notably, advancements in neuroscience have revealed similar structures in the brains of organisms such as fruit fly larvae through uncovering the entire connectome. A 2023 study [19] identified "multilayer shortcuts" within the insect brain that function similarly to the skip connections utilized in ResNets and other artificial neural networks. This finding suggests intriguing links between artificial intelligence architectures and the neural designs observed in biological systems, providing a fascinating intersection of technology and biology.

2.2. DenseNet

The success of ResNet suggested that short connections allow training deeper networks and making the model more accurate. The authors analyzed this fact and presented DenseNet (Densely Connected Convolutional Network) in 2017 [20]. They reworked the residual block and presented a dense block. The main difference from the residual block is that now the feature maps are not summed up, but concatenated, and each layer receives information from the previous ones.

The developers claim that in addition to improving efficiency, this network organization improves the flow of information and gradients throughout the network, which facilitates training, thereby helping deep networks. In addition, dense blocks have a regularizing effect that reduces overfitting in problems with small training sets.

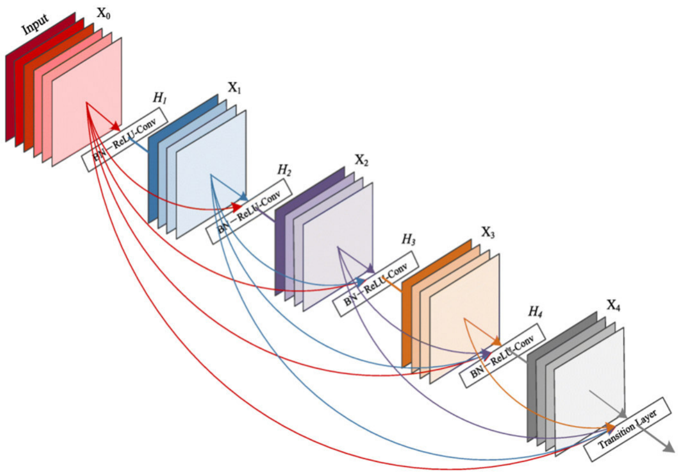


Figure 4. One of the variants for implementing the DenseNet

2.3. ResNeSt

The ResNeSt network extends the classic ResNet framework by integrating a nested skip connection strategy [21]. This approach allows for multiple pathways through which features can flow, thus enabling the model to learn from various levels of abstraction within the input data. Specifically, the architecture incorporates a series of "split-attention" blocks that enable the network to aggregate information based on the importance of different features.

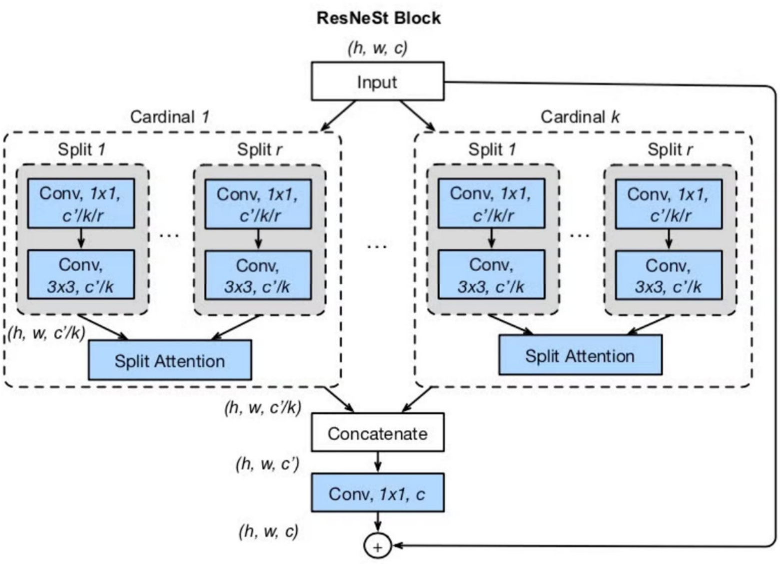


Figure 5. ResNeSt block [21]

Each block consists of two stages: the first stage applies convolutional operations to generate feature maps, while the second stage uses attention mechanisms to weigh the output of these feature maps (fig. 5). This dual-stage processing allows ResNeSt to capture both local and global contextual information, leading to enhanced representational power.

The model's ability to efficiently capture intricate patterns in images translates into significant gains in various tasks, including image classification, object detection, and semantic segmentation. It boosts the performance of downstream models such as Mask R-CNN, Cascade R-CNN and DeepLabV3.

3. Implementation Details

We used models implemented in timm (PyTorch Image Models) with Apache-2.0 licence [22]. timm is a deep-learning framework that offers a comprehensive suite of state-of-the-art computer vision models, encompassing layers, utilities, optimizers, schedulers, data handling tools, augmentation techniques, and training/validation scripts. Notably, it empowers users to replicate ImageNet training outcomes.

We have chosen a non-standard optimizer, Adan [23], which is probably not as well-known as SGD or Adam. Adan (ADAptive Nesterov momentum algorithm) introduces a Nesterov momentum estimation approach to achieve stable and precise estimations of the first and second moments of gradients in adaptive gradient algorithms, enhancing acceleration. Compared to previous adaptive optimizers like Adam, Adan demonstrates provably faster convergence [24]. Empirical results reveal that Adan delivers significantly improved generalization performance, outperforming state-of-the-art deep optimizers on both CNNs and transformers. We set the learning rate to 10^{-3} and weight decay to 10^{-4}.

We used two datasets and split each into training, validation, and test parts in a $70 : 15 : 15$ ratio. Then we combined each part of the first dataset and the second one respectively (training with training, and so on). During training, we monitored the validation balanced accuracy and saved the best model for it, which was then used on the test data. All models were trained for 25 epochs. The input data shape for each model is $3 \times 128 \times 128$ (CHW order). The batch size is set to 64 for all models. All the algorithms were trained using an NVIDIA RTX 3080 Ti (12 GB) with Python 3.11, PyTorch 2.7 and CUDA 11.8.

4. On-The-Fly Data Augmentation

Data augmentation plays a vital role in enhancing the performance of deep learning models, particularly in domains with limited data [25]. On-the-fly data augmentation (also online augmentation) refers to the process of generating augmented data in real-time during the training process, rather than pre-computing and storing augmented samples before training begins. This method has gained popularity due to its efficiency and effectiveness in reducing overfitting, improving model generalization, and increasing the diversity of the training dataset. We have chosen the well-proven Albumentations [26] to create augmentations. Training augmentations consisted of affine transform (scale, translate and rotate), H&E staining, Gaussian noise, blurring (motion, median or box blur), random changes of brightness/contrast, image compression (JPEG) and grid dropout. We perform these augmentations with some probabilities on-the-fly.

5. Weighted Random Sampling

Weighted random sampling (WRS) is a method for selecting items from a collection where each item has a different probability of being chosen, based on its assigned weight [27]. Higher weights indicate a greater chance of an item being selected.

This is in contrast to standard random sampling where each item has an equal probability of being chosen. This technique can be applied to solve the problems of imbalanced dataset. We use it during training, where the sample weight is inversely proportional to the class frequency (from the training set). This is one of the simplest techniques, which is also natively implemented in PyTorch.

6. Loss Functions

In our study, we explored several loss functions to optimize the performance of our mitotic figure classification models, specifically focusing on Focal Loss, Recall Loss, and Focal Recall Loss. Each of these loss functions provides unique advantages tailored to the challenges presented by imbalanced datasets.

6.1. Focal Loss

Focal Loss is designed to address class imbalance by applying a modulating factor to the standard cross-entropy loss [28]. This loss function is particularly effective in scenarios where the number of negative samples vastly exceeds the number of positive samples, as is common in mitotic figure identification tasks. The formulation of Focal Loss is given by:

$$L_F = -\alpha_t (1 - p_t)^\gamma \log p_t, \quad (3)$$

where p_t is the predicted probability of the true class, α_t is a balancing factor for each class, and γ is a focusing parameter that adjusts the rate at which easy examples are down-weighted

6.2. Recall Loss

Recall Loss focuses specifically on the model's ability to correctly identify positive instances in the dataset [29]. Unlike traditional loss functions that may prioritize overall accuracy, Recall Loss emphasizes the true positive rate, guiding the model toward minimizing false negatives. The formulation for Recall Loss can be defined as:

$$L_R = 1 - \frac{1}{N} \sum_{i=1}^{N} \frac{TP_i}{TP_i + FN_i}, \quad (4)$$

where TP and FN can be calculated as contiguous way (let y be the one-hot encoded target vector and $\hat{y}$ be the predicted probabilities for each class):

$$TP = \sum_{i=1}^{N} (y_i \cdot \hat{y}_i), \quad FN = \sum_{i=1}^{N} \big(y_i \cdot (1 - \hat{y}_i)\big). \quad (5)$$

The objective of Recall Loss is to ensure high sensitivity, which is critical in pathological assessments where missing atypical mitotic figures could lead to severe diagnostic errors.

6.3. Focal Recall Loss

Focal Recall Loss combines the advantages of Focal Loss and Recall Loss, targeting both the challenge of class imbalance and the importance of high sensitivity [30, 31]. This loss function modifies the standard Recall Loss by incorporating a focusing parameter that allows for downweighting contributions from well-classified examples. The formulation for Focal Recall Loss is expressed as:

$$L_{FR} = -\frac{1}{N} \sum_{i=1}^{N} \left(1 - \frac{TP_i}{TP_i + FN_i}\right)^\beta \cdot \alpha_t (1 - p_t)^\gamma \log p_t. \quad (4)$$

In this equation, pt represents the predicted probability of the true class, while the additional terms αt , γ and β serve to modulate the loss based on classification performance. Focal Recall Loss aims to enhance the model's focus on correctly identifying atypical mitotic figures in a challenging dataset, effectively balancing the need for precision in the context of clinical relevance. For all experiments, we set the values of αt , γ and β to 1, 2 and 2, respectively.

Results

In this section, we present the performance metrics of various models evaluated on the combined dataset (AMI-Br + MIDOG25) using three different loss functions: Focal Loss (FL), Recall Loss (RL) and Focal Recall Loss (FRL). The models assessed include ResNet and DenseNet architectures of various depths, specifically ResNet-34D, ResNetV2-34D, DenseNet-121, and ResNeSt-26D. Table 2 summarizes the test results, showcasing model performance.

Table 2. **Test results on the combined dataset (ami-br + midog25)**

Model	BA (%)	F_1 (%)	AUROC (%)
ResNet-34D (FL)	79.21	73.78	88.92
ResNet-34D (RL)	79.49	74.50	88.89
ResNet-34D (FRL)	79.55	74.40	88.09
ResNetV2-34D (FL)	80.72	72.83	89.89
ResNetV2-34D (RL)	**84.93**	76.73	**92.12**
ResNetV2-34D (FRL)	82.42	73.26	90.85
DenseNet-121 (FL)	81.70	73.38	89.28
DenseNet-121 (RL)	82.81	76.67	90.39
DenseNet-121 (FRL)	81.99	75.96	89.85
ResNeSt-26D (FL)	82.71	74.66	90.56
ResNeSt-26D (RL)	82.48	**76.81**	90.27
ResNeSt-26D (FRL)	81.22	70.52	89.32

Quite good quality metrics were obtained for all models. The best quality in terms of BA (84.93) and AUROC (92.12) was shown by the ResNetV2-34D model with Recall Loss. The best model in terms of F1 (76.81) was ResNeSt-26D also with Recall Loss.

We can note that in many cases, using Recall Loss or Focal Recall Loss also has a positive effect on the results. Although here, extensive experiments with more careful tuning of the optimizer, network hyperparameters or loss function may be required. It is obvious that there is a gap for improving the results, and also that the existing labelling may be inaccurate, which introduces an error.

Conclusions

In this work, we presented several ResNet-like models for normal/atypical mitotic figure classification, a relatively new task in computational histology. We tested several well-known models, such as ResNet-34D, ResNetV2-34D, DenseNet-121, and ResNeSt-26D. We used two public datasets, AMi-Br and MIDOG25, to best capture the existing variable real-world histological imaging conditions. We also considered several error functions: Focal Loss (FL), Recall Loss (RL), and Focal Recall Loss (FRL).

Good results were demonstrated, which obviously can be further improved. Several potential directions of work can be noted: (1) using other neural network architectures, such as Transformers or others, (2) additional use of other datasets, (3) conducting experiments to test generalizability to different tissues, scanners, etc.

References

1. T. A. Donovan, F. M. Moore, C. A. Bertram, R. Luong, P. Bolfa, R. Klopfleisch, H. Tvedten, E. N. Salas, D. B. Whitley, M. Aubreville, D. J. Meuten, Mitotic figuresnormal, atypical, and imposters: A guide to identification, Veterinary Pathology 58 (2020) 243257. http://dx.doi.org/10.1177/0300985820980049

2. A. Lashen, M. S. Toss, M. Alsaleem, A. R. Green, N. P. Mongan, E. Rakha, The characteristics and clinical significance of atypical mitosis in breast cancer, Modern Pathology 35 (2022) 13411348. http://dx.doi.org/10.1038/s41379-022-01080-0

3. C. A. Bertram, A. Bartel, T. A. Donovan, M. Kiupel, Atypical mitotic figures are prognostically meaningful for canine cutaneous mast cell tumors, Veterinary Sciences 11 (2023) 5. http://dx.doi.org/10.3390/vetsci11010005

4. M. Cui, D. Y. Zhang, Artificial intelligence and computational pathology, Laboratory Investigation 101 (2021) 412422. http://dx.doi.org/10.1038/s41374-020-00514-0

5. A. N. N. Wong, Z. He, K. L. Leung, C. C. K. To, C. Y. Wong, S. C. C. Wong, J. S. Yoo, C. K. R. Chan, A. Z. Chan, M. D. Lacambra, M. H. Y. Yeung, Current developments of artificial intelligence in digital pathology and its future clinical applications in gastrointestinal cancers, Cancers 14 (2022) 3780. http://dx.doi.org/10.3390/cancers14153780

6. M. Aubreville, J. Ganz, J. Ammeling, T. A. Donovan, R. H. Fick, K. Breininger, C. A. Bertram, Deep Learning-based Subtyping of Atypical and Normal Mitoses using a Hierarchical Anchor-free Object Detector, Springer Fachmedien Wiesbaden, 2023, p. 189195. http://dx.doi.org/10.1007/978-3-658-41657-7_40

7. J. Ammeling, M. Aubreville, S. Banerjee, C. A. Bertram, K. Breininger, D. Hirling, P. Horvath, N. Stathonikos, M. Veta, Mitosis domain generalization challenge 2025, 2025. https://zenodo.org/doi/10.5281/zenodo.15077361

8. C. A. Bertram, V. Weiss, T. A. Donovan, S. Banerjee, T. Conrad, J. Ammeling, R. Klopfleisch, C. Kaltenecker, M. Aubreville, Histologic Dataset of Normal and Atypical Mitotic Figures on Human Breast Cancer (AMi-Br), Springer Fachmedien Wiesbaden, 2025, p. 113118. http://dx.doi.org/10.1007/978-3-658-47422-5_25

9. S. Banerjee, V. Weiss, T. A. Donovan, R. H. J. Fick, T. Conrad, J. Ammeling, N. Porsche, R. Klopfleisch, C. Kaltenecker, K. Breininger, M. Aubreville, C. A. Bertram, Benchmarking deep learning and vision foundation models for atypical vs. normal mitosis classification with cross-dataset evaluation, 2025. https://arxiv.org/abs/2506.21444. doi:10.48550/ARXIV.2506.21444.

10. J. Ammeling, M. Aubreville, S. Banerjee, C. A. Bertram, K. Breininger, D. Hirling, P. Horvath, N. Stathonikos, M. Veta, Mitosis domain generalization challenge 2025, 2025. https://zenodo.org/doi/10.5281/zenodo.15077360

11. M. Veta, Y. J. Heng, N. Stathonikos, B. E. Bejnordi, F. Beca, T. Wollmann, K. Rohr, M. A. Shah, D. Wang, M. Rousson, M. Hedlund, D. Tellez, F. Ciompi, E. Zerhouni, D. Lanyi, M. Viana, V. Kovalev, V. Liauchuk, H. A. Phoulady, T. Qaiser, S. Graham, N. Rajpoot, E. Sj¨oblom, J. Molin, K. Paeng, S. Hwang, S. Park, Z. Jia, E. I.-C. Chang, Y. Xu, A. H. Beck, P. J. van Diest, J. P. Pluim, Predicting breast tumor proliferation from wholeslide images: The tupac16 challenge, Medical Image Analysis 54 (2019) 111121. http://dx.doi.org/10.1016/j.media.2019.02.012

12. M. Aubreville, N. Stathonikos, C. A. Bertram, R. Klopfleisch, N. ter Hoeve, F. Ciompi, F. Wilm, C. Marzahl, T. A. Donovan, A. Maier, J. Breen, N. Ravikumar, Y. Chung, J. Park, R. Nateghi, F. Pourakpour, R. H. Fick, S. Ben Hadj, M. Jahanifar, A. Shephard, J. Dexl, T. Wittenberg, S. Kondo, M. W. Lafarge, V. H. Koelzer, J. Liang, Y. Wang, X. Long, J. Liu, S. Razavi, A. Khademi, S. Yang, X. Wang, R. Erber, A. Klang, K. Lipnik, P. Bolfa, M. J. Dark, G. Wasinger,

M. Veta, K. Breininger, Mitosis domain generalization in histopathology images the midog challenge, Medical Image Analysis 84 (2023) 102699. http://dx.doi.org/10.1016/j.media.2022.102699

13. J. Mauricio, I. Domingues, J. Bernardino, Comparing vision transformers and convolutional neural networks for image classification: A literature review, Applied Sciences 13 (2023) 5521. http://dx.doi.org/10.3390/app13095521
14. S. Basodi, C. Ji, H. Zhang, Y. Pan, Gradient amplification: An efficient way to train deep neural networks, Big Data Mining and Analytics 3 (2020) 196207. http://dx.doi.org/10.26599/BDMA.2020.9020004
15. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016, p. 770778. http://dx.doi.org/10.1109/CVPR.2016.90
16. K. He, X. Zhang, S. Ren, J. Sun, Identity Mappings in Deep Residual Networks, Springer International Publishing, 2016, p. 630645. http://dx.doi.org/10.1007/978-3-319-46493-0_38
17. H. Li, Z. Xu, G. Taylor, C. Studer, T. Goldstein, Visualizing the loss landscape of neural nets, in: Proceedings of the 32nd International Conference on Neural Information Processing Systems, NIPS'18, Curran Associates Inc., Red Hook, NY, USA, 2018, p. 63916401.
18. L. Wang, B. Shen, N. Zhao, Z. Zhang, Is the skip connection provable to reform the neural network loss landscape?, in: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI'20, 2021.
19. M. Winding, B. D. Pedigo, C. L. Barnes, H. G. Patsolic, Y. Park, T. Kazimiers, A. Fushiki, I. V. Andrade, A. Khandelwal, J. Valdes-Aleman, F. Li, N. Randel, E. Barsotti, A. Correia, R. D. Fetter, V. Hartenstein, C. E. Priebe, J. T. Vogelstein, A. Cardona, M. Zlatic, The connectome of an insect brain, Science 379 (2023). http://dx.doi.org/10.1126/science. add9330
20. G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017. http://dx.doi.org/10.1109/CVPR.2017.243
21. H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Lin, Z. Zhang, Y. Sun, T. He, J. Mueller, R. Manmatha, M. Li, A. Smola, Resnest: Split-attention networks, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2022, p. 27352745. http://dx.doi.org/10.1109/CVPRW56347.2022.00309
22. R. Wightman, Pytorch image models, https://github.com/rwightman/pytorch-image-models, 2019. doi:10.5281/zenodo.4414861.
23. X. Xie, P. Zhou, H. Li, Z. Lin, S. Yan, Adan: Adaptive nesterov momentum algorithm for faster optimizing deep models, IEEE Transactions on Pattern Analysis and Machine Intelligence 46 (2024) 95089520. http://dx.doi.org/10.1109/TPAMI.2024.3423382
24. A. R. Ismail, M. Z. R. Azhary, N. A. Hitam, Evaluating Adan vs. Adam: An Analysis of Optimizer Performance in Deep Learning, Springer Nature Switzerland, 2025, p. 251263. http://dx.doi.org/10.1007/978-3-031-82931-4_19.
25. J. Ma, C. Hu, P. Zhou, F. Jin, X. Wang, H. Huang, Review of image augmentation used in deep learning-based material microscopic image segmentation, Applied Sciences 13 (2023) 6478. http://dx.doi.org/10.3390/app13116478
26. A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, A. A. Kalinin, Albumentations: Fast and flexible image augmentations, Information 11 (2020) 125. http://dx.doi.org/10.3390/info11020125.
27. P. Efraimidis, P. Spirakis, Weighted Random Sampling: 2005; Efraimidis, Spirakis, Springer US, 2008, p. 10241027. http://dx.doi.org/10.1007/978-0-387-30162-4_478
28. T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, Focal loss for dense object detection, IEEE Transactions on Pattern Analysis and Machine Intelligence 42 (2020) 318327. http://dx.doi.org/10.1109/TPAMI.2018.2858826
29. J. Tian, N. C. Mithun, Z. Seymour, H.-P. Chiu, Z. Kira, Striking the right balance: Recall loss for semantic segmentation, in: 2022 International Conference on Robotics and Automation (ICRA), IEEE, 2022, p. 50635069. http://dx.doi.org/10.1109/ICRA46639.2022.9811702
30. Z. Chen, Z. Wang, J. Qiu, The alzheimers disease segmentation model based on focal recall loss, Computational Science and Applications 12 (2022) 95–107. doi:10.12677/CSA.2022. 121011.
31. X. Liu, L. Wang, L. Ma, C. Wang, Drfl: Dynamic-recall focal loss for image classification and segmentation, Applied Artificial Intelligence 38 (2024). http://dx.doi.org/10.1080/ 08839514.2024.2411845.