



Article

Deep Learning Approach for Differentiating Etiologies of Pediatric Retinal Hemorrhages: A Multicenter Study

Pooya Khosravi ^{1,2,3} , Nolan A. Huck ^{1,2}, Kouros Shahraki ^{1,2} , Stephen C. Hunter ⁴, Clifford Neil Danza ^{1,2}, So Young Kim ⁵, Brian J. Forbes ⁶, Shuan Dai ⁷, Alex V. Levin ⁸, Gil Binenbaum ⁶, Peter D. Chang ^{3,9} and Donny W. Suh ^{1,2,*}

- ¹ Department of Ophthalmology, School of Medicine, University of California, Irvine, CA 92697, USA; pooyak@hs.uci.edu (P.K.); nahuck@hs.uci.edu (N.A.H.); kouros.shahyar@gmail.com (K.S.); cdanza@hs.uci.edu (C.N.D.)
- ² Gavin Herbert Eye Institute, University of California, Irvine, CA 92697, USA
- ³ Donald Bren School of Information and Computer Sciences, University of California, Irvine, CA 92697, USA; changp6@hs.uci.edu
- ⁴ School of Medicine, University of California, 900 University Ave, Riverside, CA 92521, USA; stephen.hunter@medsch.ucr.edu
- ⁵ Department of Ophthalmology, College of Medicine, Soonchunhyang University, Cheonan 31151, Chungcheongnam-do, Republic of Korea; ophdrkim@gmail.com
- ⁶ Division of Ophthalmology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; forbesb@email.chop.edu (B.J.F.); binenbaum@chop.edu (G.B.)
- ⁷ Department of Ophthalmology, Queensland Children's Hospital, South Brisbane, QLD 4101, Australia; shuan.dai@health.qld.gov.au
- ⁸ Department of Ophthalmology, Flaum Eye Institute, Golisano Children's Hospital, Rochester, NY 14642, USA; alex_levin@urmc.rochester.edu
- ⁹ Department of Radiological Sciences, School of Medicine, University of California, Irvine, CA 92697, USA
- * Correspondence: donnys@hs.uci.edu; Tel.: +1-949-824-4122; Fax: +1-949-824-4015



Citation: Khosravi, P.; Huck, N.A.; Shahraki, K.; Hunter, S.C.; Danza, C.N.; Kim, S.Y.; Forbes, B.J.; Dai, S.; Levin, A.V.; Binenbaum, G.; et al. Deep Learning Approach for Differentiating Etiologies of Pediatric Retinal Hemorrhages: A Multicenter Study. *Int. J. Mol. Sci.* **2023**, *24*, 15105. <https://doi.org/10.3390/ijms242015105>

Academic Editor: Hao Lin

Received: 25 August 2023

Revised: 29 September 2023

Accepted: 10 October 2023

Published: 12 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Retinal hemorrhages in pediatric patients can be a diagnostic challenge for ophthalmologists. These hemorrhages can occur due to various underlying etiologies, including abusive head trauma, accidental trauma, and medical conditions. Accurate identification of the etiology is crucial for appropriate management and legal considerations. In recent years, deep learning techniques have shown promise in assisting healthcare professionals in making more accurate and timely diagnosis of a variety of disorders. We explore the potential of deep learning approaches for differentiating etiologies of pediatric retinal hemorrhages. Our study, which spanned multiple centers, analyzed 898 images, resulting in a final dataset of 597 retinal hemorrhage fundus photos categorized into medical (49.9%) and trauma (50.1%) etiologies. Deep learning models, specifically those based on ResNet and transformer architectures, were applied; FastViT-SA12, a hybrid transformer model, achieved the highest accuracy (90.55%) and area under the receiver operating characteristic curve (AUC) of 90.55%, while ResNet18 secured the highest sensitivity value (96.77%) on an independent test dataset. The study highlighted areas for optimization in artificial intelligence (AI) models specifically for pediatric retinal hemorrhages. While AI proves valuable in diagnosing these hemorrhages, the expertise of medical professionals remains irreplaceable. Collaborative efforts between AI specialists and pediatric ophthalmologists are crucial to fully harness AI's potential in diagnosing etiologies of pediatric retinal hemorrhages.

Keywords: artificial intelligence; deep learning; pediatrics; retinal hemorrhage

1. Introduction

Retinal hemorrhages can pose substantial diagnostic challenges due to their association with various systemic and ocular diseases [1]. The retina comprises a total of 10 layers, with the outermost layer being the retinal pigmented epithelium (RPE), followed by 9 other

layers collectively known as the neurosensory retina. The innermost layer is referred to as the inner limiting membrane (ILM) [2]. Hemorrhages can occur within the layers of the retina (intraretinal) or above the retina itself (preretinal). Subretinal hemorrhages are an accumulation of blood between the neurosensory retina and the RPE. If the blood extends into the vitreous humor, it is referred to as vitreous hemorrhage. Intraretinal hemorrhages that occur in the superficial layers of the retina exhibit a linear streaking appearance, following the pattern of the nerve fibers, often referred to as flame or splinter hemorrhages due to their distinct appearance. On the other hand, intraretinal hemorrhages that occur in the deeper layers of the retina have a more rounded shape and are referred to as dot or blot hemorrhages, depending on their size [2–4].

The differential diagnosis of retinal hemorrhages is extensive and includes [3,5] coagulopathy, blood dyscrasia, severe anemia, hemolytic uremic syndrome, endocarditis, vasculitis [2,3], normal birth [6], raised intracranial pressure (ICP), glutaric aciduria type 1, meningitis and Terson syndrome [7]. Retinal hemorrhages can also be diagnostic indicators for severe conditions such as abusive head trauma [1,2,8–11]. Common causes of retinal hemorrhage in adults include diabetic retinopathy, hypertensive retinopathy, retinal vein occlusion [2,3,12], trauma [9,13], and hematological disorders [2].

Fundus photography increases our ability to visualize, document, and monitor retinal hemorrhages [14,15]. While the interpretation of these images heavily relies on the trained eye of an experienced ophthalmologist, there is potential for further improving interpretation and consistency. This is where artificial intelligence (AI) and machine learning (ML) technologies can enhance ophthalmologic practice, offering advanced computational tools that complement, rather than replace, the expert judgment of clinical professionals. This helps reduce variability and ensure a more standardized approach across observers [15]. By doing so, AI and ML contribute to the ongoing development of improved diagnostic capabilities in ophthalmology.

Previous research has demonstrated the use of AI and ML to analyze fundus photographs [14–16]. These papers primarily focused on disease detection [17,18], classification and grading [18,19], segmentation [20,21] and prediction [22] for retinopathy of prematurity (ROP) [23,24], diabetic retinopathy [18,25,26], age-related macular degeneration [21,27–30], glaucoma [31–33], cataracts [34], lacrimal disorders [35], keratoconus [36], amblyopia [37] and optic nerve diseases [38,39]. Studies have demonstrated that AI and ML algorithms, including deep learning models such as convolutional neural networks (CNNs), RNNs, and Transformers, can accurately identify and classify different types of retinal lesions, such as hemorrhages, exudates, and micro-aneurysms [25,26]. These algorithms can analyze large datasets of fundus photographs and provide automated assessments of disease severity and progression. Furthermore, AI and ML have been utilized to predict the risk of developing certain retinal diseases, such as geographic atrophy, based on fundus image analysis [27,29]. By training algorithms on large datasets of fundus photographs and corresponding patient data, these technologies can identify patterns and markers that are indicative of disease progression or future complications [29].

Unlike previous studies that primarily aimed to describe abnormal exams, our study seeks to develop and validate CNN and transformer models to differentiate between patients with known retinal hemorrhage. This is an especially challenging task as retinal hemorrhages can be difficult to interpret even for trained experts. To date, despite the progress made in applying AI in ophthalmology, there is still a lack of research exploring the potential of computational tools in differentiating the etiologies of retinal hemorrhages.

2. Results

Our study spanned multiple centers and involved a comprehensive analysis of 898 retinal hemorrhage (RH) fundus images of which 301 images were excluded due to the absence of retinal hemorrhage, resulting in a final dataset comprising 597 images. The images were divided into medical ($n = 298$, 49.9%), and trauma etiologies ($n = 299$, 50.1%). The detailed

distribution of the number of photos corresponding to each diagnosis is provided in Table 1. Figure 1 showcases representative examples of these diagnoses.

Table 1. Dataset diagnosis characteristics.

| Diagnosis | All (<i>n</i> = 599) | Training (<i>n</i> = 343) | Validation (<i>n</i> = 127) | Test (<i>n</i> = 127) |
|--------------------------|-----------------------|----------------------------|------------------------------|------------------------|
| Medical | 298 | 172 | 61 | 65 |
| Retinal Vascular Disease | 102 | 54 | 20 | 27 |
| Leukemia | 89 | 46 | 16 | 17 |
| Papilledema | 19 | 12 | 4 | 3 |
| Coagulopathy | 88 | 60 | 22 | 18 |
| Trauma | 299 | 171 | 66 | 62 |
| Accidental Trauma | 18 | 11 | 3 | 4 |
| Birth Trauma | 118 | 67 | 25 | 26 |
| Abusive Head Trauma | 163 | 93 | 38 | 32 |

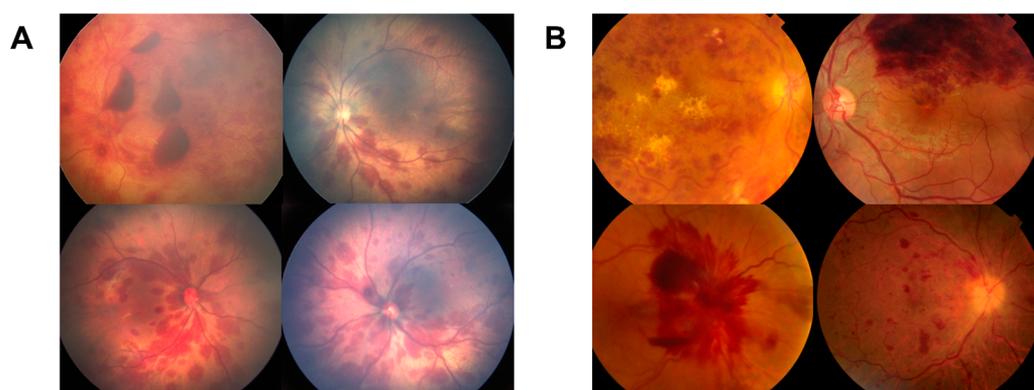


Figure 1. Representative fundus photographs from our pediatric dataset, categorized by diagnosis. This diverse collection includes images dichotomized into trauma cases (A) that encompass vaginal birth trauma (right), and abusive head trauma (left), and medical cases (B), such as ischemic CRVO (top left), coagulopathy (bottom left), BRVO (top right), and CRVO (bottom right).

Upon completing the pre-processing stage, the dataset was divided at the patient level into distinct subsets for training, validation, and testing, representing 60%, 20%, and 20% of the total dataset, respectively. The training subset (*n* = 343; 172 medical, 171 trauma) was used to adapt the model's parameters and to learn the underlying patterns in differentiating medical from traumatic retinal hemorrhages. The validation subset (*n* = 127; 61 medical, 66 trauma) played a crucial role in tuning hyperparameters and selecting the best-performing model, ensuring the chosen model was robust yet not overfitted to the training data. The testing subset (*n* = 127; 65 medical, 62 trauma) was then used to provide an unbiased evaluation of the final model's performance in unseen data. Dividing the data at the patient level ensured that all images from the same patient were contained within the same subset, minimizing the risk of data leakage between the subsets and thus providing a more rigorous and trustworthy evaluation of the model's performance. This approach aligns with best practices in machine learning and provides a sound basis for extrapolating the findings to a broader population. The performance metrics of the models based on the ResNet architecture were highly compelling. Among the ResNet models, the ResNet18 model exhibited the most remarkable results, with an AUC of 0.9506 and an accuracy of 88.98%. It also demonstrated the highest sensitivity of 96.77% and a PPV of 84.29%. The ResNet101 model achieved an AUC of 0.9449, while attaining the highest accuracy, specificity, and PPV of 89.76%, 90.77%, and 90.16%, respectively out of the five models. Comprehensive performance metrics are presented in Table 2. AUC graphs (Figures S1 and S2) and confusion matrices (Figure S3) are included in the supplementary materials.

Table 2. Performance metrics of the models.

| Model | Accuracy | AUC | Specificity | Sensitivity | PPV | NPV |
|--------------|---------------|---------------|---------------|---------------|---------------|---------------|
| ResNet18 | 88.98% | 0.9506 | 83.08% | 96.77% | 84.29% | 94.74% |
| ResNet34 | 86.61% | 0.9437 | 87.69% | 87.10% | 86.89% | 86.36% |
| ResNet50 | 87.40% | 0.9467 | 84.62% | 91.94% | 84.85% | 90.16% |
| ResNet101 | 89.76% | 0.9449 | 90.77% | 90.32% | 90.16% | 89.39% |
| ResNet152 | 88.19% | 0.9365 | 84.62% | 93.55% | 85.07% | 91.67% |
| ResAttNet56 | 87.40% | 0.9400 | 83.08% | 93.55% | 83.82% | 91.53% |
| ViT-Small | 79.53% | 0.8945 | 78.46% | 82.26% | 78.12% | 80.95% |
| FastViT-SA12 | 90.55% | 0.9628 | 96.92% | 85.48% | 96.30% | 86.30% |
| FastViT-SA24 | 88.19% | 0.9462 | 87.69% | 90.32% | 87.30% | 89.06% |

AUC: Area under the receiver operating characteristic curve, PPV: Positive predictive value, NPV: Negative predictive value. Specificity, Sensitivity, PPV, and NPV were calculated at the Youden Index. Bolded values indicate the highest performance among the models.

The FastViT-SA12 model with the highest accuracy misclassified 2 medical and 10 trauma cases, while the ResNet18 model with the highest sensitivity misclassified 8 medical and 6 trauma cases. ViT-Small with the lowest accuracy misclassified 12 medical and 14 trauma cases. The misclassifications by diagnosis are presented in Tables S1 and S2. It should be noted that the errors made by all models exhibited a discernible pattern. Specifically, cases of acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL), as well as other hematologic disorders and coagulopathies that caused significant retinal hemorrhaging, were often misclassified by all models as traumatic. This confusion likely arose due to the similarity in the number and extent of hemorrhaging between these medical conditions and trauma-related cases. Conversely, two instances of accidental and abusive head trauma, characterized by minimal retinal hemorrhage, were incorrectly identified as medical in origin. On the other hand, cases of papilledema and birth trauma were all classified correctly. Examples of these misclassifications are provided in Figure 2.

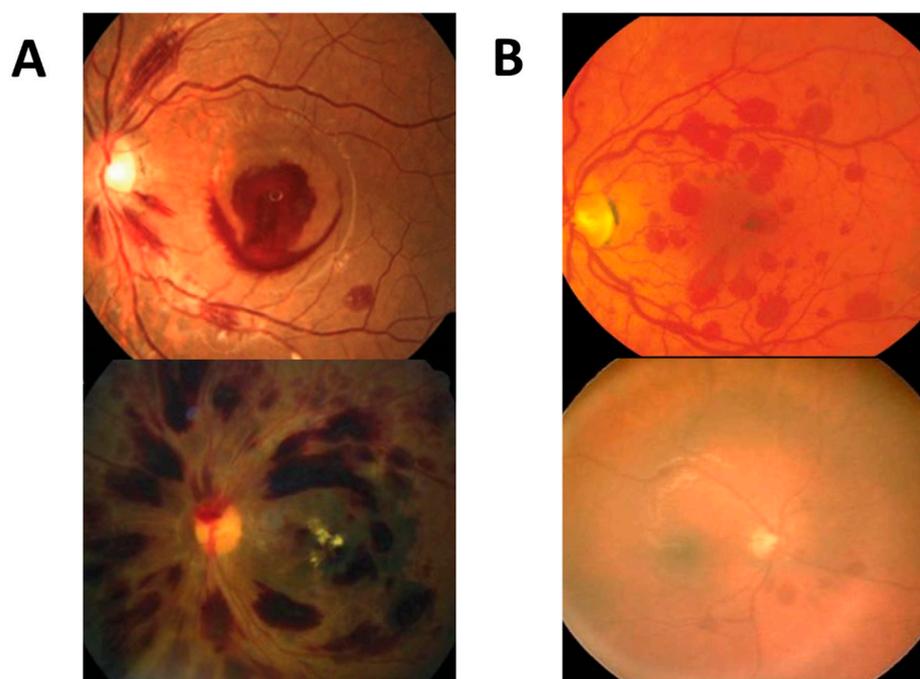


Figure 2. The two images on the left (A) illustrate medical cases (top left: leukemia, bottom left: aplastic anemia) misclassified as trauma, and the two on the right (B) show traumatic cases (top right: abusive head trauma (AHT), bottom right: accidental trauma) wrongly predicted as medical.

3. Discussion

The findings of this study illustrate the potential of deep learning models, particularly transformer models with attention mechanisms, in differentiating medical versus traumatic pediatric retinal hemorrhages (RH). This study is unique in that it focuses specifically on differentiating traumatic from medical retinal hemorrhages, rather than comparing normal and abnormal images. This distinction has significant medicolegal implications, particularly for cases involving abusive head trauma.

Although this is the first study that investigates distinguishing etiologies of retinal hemorrhages from fundus photos, there have been previous studies focused on detecting the presence, segmentation, or identifying the layer of retinal hemorrhages from fundus images [40–43].

The superior performance of FastVit-SA12 and ResNet18 in our experiments, as highlighted by the two highest AUCs among the evaluated models, underscores the importance of carefully selecting model architecture in accordance with the size of the available dataset. Both ResNet18's and FastVit-SA12's relatively streamlined architectures demonstrated a notable advantage in adapting to our limited dataset, potentially due to their smaller parameter count compared to the more intricate designs of larger ResNet and transformer models (Table 3). The relatively fewer parameters in FastVit-SA12 and ResNet18, in comparison to the larger models, may have conferred an advantage in curtailing overfitting on the limited data, consequently enhancing their ability to generalize effectively on both the validation set and the test set. Models characterized by heightened complexity, such as ResNet101, ResNet152, ViT-Small, and FastVit-SA24, while beneficial for larger datasets, might exhibit diminished efficiency on smaller datasets and increased susceptibility to overfitting to the training data. This aligns with the principle of Occam's razor in model selection, which suggests that simpler models are often preferable when they perform as well, or better than, more complex models on the same task [44].

Table 3. Model characteristics of the models.

| Model | Number of Layers | Trainable Parameters (in Millions) |
|--------------|------------------|------------------------------------|
| ResNet18 | 18 | 11.4 |
| ResNet34 | 32 | 21.5 |
| ResNet50 | 50 | 24.0 |
| ResNet101 | 101 | 43.1 |
| ResNet152 | 152 | 58.7 |
| ResAttNet56 | 56 | 29.8 |
| ViT-Small | 12 | 22.5 |
| FastVit-SA12 | 12 | 10.5 |
| FastVit-SA24 | 24 | 20.5 |

These results highlight the potential of transformers and CNNs as robust tools for distinguishing the medical versus traumatic RH in pediatric patients, even in scenarios involving relatively limited datasets. This finding is a testament to the need for striking a balance between model complexity and the risk of overfitting when working with limited datasets. Additional research using larger datasets would be advantageous in validating these findings and exploring the performance of more complex models.

The observed patterns of misclassification within our study provide valuable insights into areas for model optimization. Instances where extensive hemorrhages due to leukemia were identified as trauma, and a few accidental head trauma cases were predicted as medical. With an extensive RH, it may hide the underlying cotton wool spots, hard exudates that may help AI to make the correct diagnosis of medical induced RH. Also, if mild traumas are present in localized areas, AI may mistakenly interpret the image as sectoral RH. These findings present an opportunity for targeted refinements that could effectively enhance model performance. Furthermore, these subtle distinctions between medical and traumatic cases offer an avenue to deepen our comprehension of various pathophysiological

presentations, potentially leading to the integration of supplementary clinical features or fine-tuned training strategies. Overall, the results accentuate the potential of the model to be a valuable diagnostic tool for differentiating retinal hemorrhages, while continually striving for enhancement and adaptation to the multifaceted nature of these conditions.

While our study aims to explore the potential of computational tools in differentiating the etiologies of retinal hemorrhages, it is important to acknowledge the expertise of experienced ophthalmologists as a benchmark in this field. The goal is to complement their expertise with AI and ML technologies, rather than replace it.

Despite the encouraging findings, our study is not without limitations. While our dataset encompasses a diverse range of RH etiologies, it is relatively small and constrained by factors such as photo quality, reliance on expert judgment, and historical chart information. The inconsistency in technologies employed to capture the images, coupled with the lack of standardization in luminance, magnification, and field, further contributes to these limitations. Additionally, some cases presented with mixed diagnoses, such as increased intracranial pressure with optic disc edema and trauma, adding complexity to classification. Our study also acted by proxy, with diagnoses predetermined based on medical history, laboratory and radiologic findings, physical examination, and multidisciplinary assessments. To address these limitations in future research, efforts should focus on acquiring a larger, higher-quality dataset that encompasses a more diverse set of retinal hemorrhage etiologies, while adhering to standardized image acquisition protocols to ensure consistency and reliability.

We conclude AI technology may be a helpful tool for assisting medical professionals in accurately distinguishing between medical and trauma-induced retinal hemorrhages. Of course, careful attention must always be paid to medical history, laboratory and radiologic findings, physical examination, and multidisciplinary assessments.

AI systems might not determine the precise biomechanism and degree of force responsible for the trauma that induced the RH. In cases where distinguishing between abusive and accidental physical trauma is crucial for legal or investigational purposes, the expertise of skilled medical professionals remains essential [10,11].

As the field of AI in medicine continues to advance, further research and refinement of algorithms may improve its capacity to gather additional contextual information and aid in unraveling the specific details surrounding traumatic events. Collaborations between AI experts and medical professionals will be pivotal in obtaining the full potential of AI while ensuring responsible and ethical implementation in the domain of medical diagnosis. Despite its current limitations, AI remains a promising and invaluable tool in the realm of healthcare, contributing significantly to improved diagnostic accuracy and patient care.

4. Methods and Materials

4.1. Study Population

Pediatric fundus photographs were collected from Soonchunhyang University Cheonan Hospital in Korea, Children's Hospital of Philadelphia, Wills Eye Hospital, The Hospital for Sick Children (Toronto), Golisano Children's Hospital and Flaum Eye Institute in Rochester, New York, NY, USA, and Queensland Children's Hospital in Brisbane, Australia. The data collection period spanned from 2015 to 2023. These images were captured using either handheld portable devices (e.g., RetCam (RetCam 3, Natus Medical Inc., Pleasanton, CA, USA), ICON (Phoenix Clinical, Inc., Pleasanton, CA, USA), or Nonmyd7 (Kowa, Torrance, CA, USA)) or by upright standard fundus photography. We included cases that were confirmed abusive head trauma (AHT), a determination that was made through multidisciplinary child abuse team evaluation. We also included medically validated cases that were correlated with laboratory results as well as cases of accidental trauma as validated by witness, scene investigation and/or consistent physical findings. We excluded cases with uncertain etiology. The utilization of unrecognizable, anonymous images negated the requirement for Institutional Review Board (IRB) approval, given the absence of any identi-

fiable personal information associated with the participants. Photographs were provided from the personal collections of the authors.

4.2. Annotation

Image labels were assigned based on the underlying case diagnosis. These labels were categorized as either traumatic or medical conditions based on medical records from the input of multi-specialists including the child abuse experts. The determination of hemorrhage presence or absence was entrusted to a panel of pediatric ophthalmologists (D.W.S., G.B., A.V.L., B.J.F.). In situations where there were varying interpretations of the image, a consensus was achieved through thoughtful deliberation until agreement was reached on the final classification.

4.3. Data Preprocessing

Prior to analysis, all images were subjected to standardized preprocessing procedures in order to ensure uniformity and enhance variability across images from different sources and devices. This involved resizing all images to a consistent shape of 256×256 pixels using bilinear interpolation, with zero-padding as necessary. Bilinear interpolation provided a reliable method for resizing images without introducing significant artifacts or distortions. Additionally, each image was normalized per channel based on the mean and standard deviation derived from the ImageNet training set [45]. During the training process, the augmentation steps included resizing the images to 224×224 to use transfer learning, applying random contrast ranging from 0.5 to 1.5, introducing Gaussian noise with a kernel size of 11×11 and sigma ranging from 0.1 to 2, randomly flipping the images horizontally with a 50% probability, and applying cropping with a scale range of 0.08 to 1. A minimum scale range of 0.08 ensures that even smaller details within the images are considered during training, potentially capturing subtle patterns that may be diagnostically relevant. Furthermore, an additional augmentation technique known as cutout [46], involving the placement of 4 randomly positioned obscuring boxes each measuring 44×44 (Figure 3), was employed. The comprehensive combination of these data augmentation methods aimed to prevent overfitting and promote the generalizability of the dataset. During the validation and testing phases, the images were resized to a matrix shape of 224×224 and underwent per-channel normalization, mirroring the preprocessing steps implemented during training.

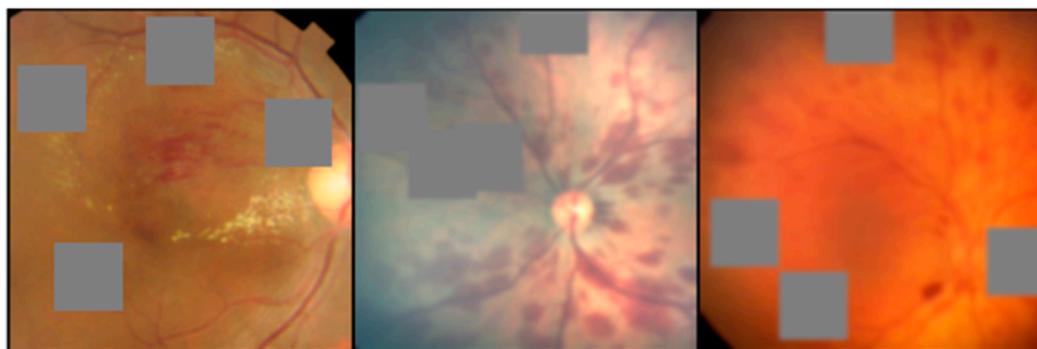


Figure 3. Depiction of the ‘cutout’ technique, which obscures random square regions to enhance the model’s focus diversity.

4.4. Algorithm Development

In this study, our focus was the training of CNN and Transformer architectures to effectively discern between retinal hemorrhages of traumatic and medical origins. To achieve this, all experiments are derived from standard ResNet-based architectures [47] (ResNet34 shown in Figure 4), Residual Attention Networks (ResAttNet56) [48] using attention residual learning, a vision Transformer (ViT) image classification model [49], and

FastViT [50], a hybrid vision transformer architecture (Table 3). Each model was fine-tuned after initialization with pretrained ImageNet weights from the PyTorch repository (torchvision 0.15.0; IMAGENET1K_V1) [51] for ResNet models and PyTorch Image Models [52,53] for ResAttNet56, ViT-small, and FastViT models.

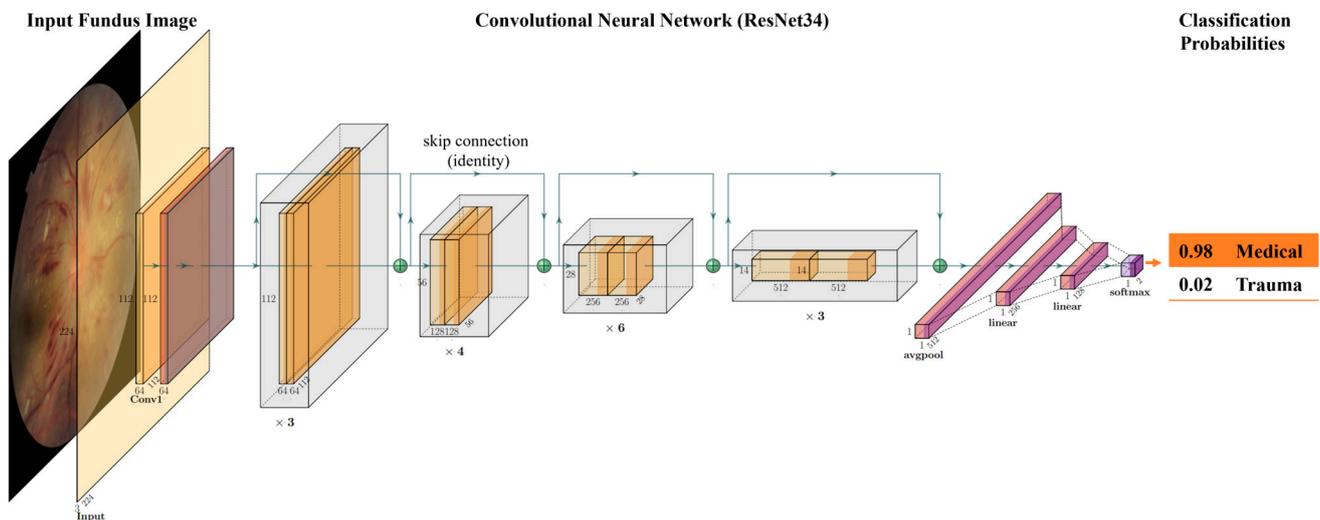


Figure 4. Visualization of the ResNet convolutional neural network (ResNet34) architecture and its application in our study. The flow diagram illustrates the journey from input fundus photos through the complex, multi-layered structure of the ResNet model, ending with the final task output.

For model fine-tuning, the default last fully connected layer is replaced with two custom fully connected hidden layers before projection to a two-element SoftMax activated logit score. The fully connected hidden layers are of dimensions 256 and 128, respectively, and employ a Rectified Linear Unit (ReLU) activation function. Additionally, the fully connected hidden layers are each implemented with batch normalization prior to ReLU activation and a dropout rate of 50%. During the fine-tuning phase, all layers of the network were unfrozen. This strategic approach enhances our models' adaptability and enables them to effectively accommodate the unique characteristics of our dataset.

Each model was trained for a total of 50 epochs, except for ViT-Small, which was trained for 250 epochs. The model was optimized using Adam optimizer [54] with a standard binary cross-entropy loss function. We employed Hyperopt [55] for hyperparameter tuning and determined that a learning rate of 0.001 and a batch size of 32 consistently yielded the best performance across all ResNet and transformer models used in our study. Optimal model selection was determined based on validation data performance. Only after this stringent evaluation did we set forth to apply the chosen model to the test dataset for final assessment.

4.5. Statistics

For the purpose of this research, the dataset was randomly partitioned into training, validation, and test cohorts at a ratio of 60:20:20, respectively, on a per-patient basis (i.e., all images from the same patient were used in the same cohort). Rigorous evaluation of the model's performance was conducted using the area under the receiver operating characteristic curve (AUC) on the independent test dataset. Furthermore, sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV) were calculated at the Youden Index threshold, which effectively distinguishes cases with traumatic causes from medical causes. The Youden Index facilitates a comprehensive assessment of the ROC curve by pinpointing an optimal cutoff value that maximizes the difference between the true positive rate (sensitivity) and the false positive rate (1-specificity).

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms242015105/s1>.

Author Contributions: Methodology, P.K., N.A.H. and P.D.C.; software, P.K., N.A.H. and P.D.C.; validation, P.K. and P.D.C.; formal analysis, P.K., N.A.H. and P.D.C.; investigation, P.K., N.A.H., S.C.H., C.N.D., S.Y.K., B.J.F., S.D., A.V.L. and G.B.; data curation, P.K., N.A.H., S.C.H., C.N.D., S.Y.K., B.J.F., S.D., A.V.L. and G.B.; writing—original draft, P.K. and K.S.; writing—review and editing, P.K., K.S., N.A.H., S.C.H., C.N.D., S.Y.K., B.J.F., S.D., A.V.L., G.B., P.D.C. and D.W.S.; visualization, P.K., K.S., N.A.H. and C.N.D.; supervision, D.W.S.; project administration, D.W.S. All authors have read and agreed to the published version of the manuscript.

Funding: D.W.S. and K.S. acknowledge an unrestricted grant from Research to Prevent Blindness to the Gavin Herbert Eye Institute at the University of California, Irvine. A.V.L. was supported in part by an unrestricted grant from Research to Prevent Blindness to the Department of Ophthalmology at the University of Rochester.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to the sensitive nature of the images, which include fundus photos with retinal hemorrhages from Abusive Head Trauma patients.

Conflicts of Interest: A.V.L. has testified as an expert on behalf of the prosecution and defense in criminal and civil matters related to the presence or absence of retinal hemorrhage. Other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Thau, A.; Saffren, B.; Zakrzewski, H.; Anderst, J.D.; Carpenter, S.L.; Levin, A. Retinal hemorrhage and bleeding disorders in children: A review. *Child Abus. Negl.* **2021**, *112*, 104901. [[CrossRef](#)] [[PubMed](#)]
2. Di Fazio, N.; Delogu, G.; Morena, D.; Cipolloni, L.; Scopetti, M.; Mazzilli, S.; Frati, P.; Fineschi, V. New Insights into the Diagnosis and Age Determination of Retinal Hemorrhages from Abusive Head Trauma: A Systematic Review. *Diagnostics* **2023**, *13*, 1722. [[CrossRef](#)] [[PubMed](#)]
3. Levin, A.V. Retinal hemorrhage: Science versus speculation. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* **2016**, *20*, 93–95. [[CrossRef](#)] [[PubMed](#)]
4. Yiu, G.; Mahmoud, T.H. Subretinal Hemorrhage. *Dev. Ophthalmol.* **2014**, *54*, 213–222. [[CrossRef](#)]
5. Forbes, B.J.; Rubin, S.E.; Margolin, E.; Levin, A.V. Evaluation and management of retinal hemorrhages in infants with and without abusive head trauma. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* **2010**, *14*, 267–273. [[CrossRef](#)]
6. Watts, P.; Maguire, S.; Kwok, T.; Talabani, B.; Mann, M.; Wiener, J.; Lawson, Z.; Kemp, A. Newborn retinal hemorrhages: A systematic review. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* **2013**, *17*, 70–78. [[CrossRef](#)] [[PubMed](#)]
7. Moshfeghi, D.M. Terson Syndrome in a Healthy Term Infant: Delivery-Associated Retinopathy and Intracranial Hemorrhage. *Ophthalmic Surg. Lasers Imaging Retin.* **2018**, *49*, e154–e156. [[CrossRef](#)] [[PubMed](#)]
8. Bechtel, K.; Stoessel, K.; Leventhal, J.M.; Ogle, E.; Teague, B.; Laviates, S.; Banyas, B.; Allen, K.; Dziura, J.; Duncan, C. Characteristics That Distinguish Accidental from Abusive Injury in Hospitalized Young Children with Head Trauma. *Pediatrics* **2004**, *114*, 165–168. [[CrossRef](#)] [[PubMed](#)]
9. Song, H.H.; Thoreson, W.B.; Dong, P.; Shokrollahi, Y.; Gu, L.; Suh, D.W. Exploring the Vitreoretinal Interface: A Key Instigator of Unique Retinal Hemorrhage Patterns in Pediatric Head Trauma. *Korean J. Ophthalmol.* **2022**, *36*, 253–263. [[CrossRef](#)] [[PubMed](#)]
10. Morad, Y.; Wygnansky-Jaffe, T.; Levin, A.V. Retinal haemorrhage in abusive head trauma. *Clin. Exp. Ophthalmol.* **2010**, *38*, 514–520. [[CrossRef](#)] [[PubMed](#)]
11. Togioka, B.M.; Arnold, M.A.; Bathurst, M.A.; Ziegfeld, S.M.; Nabaweesei, R.; Colombani, P.M.; Chang, D.C.; Abdullah, F. Retinal Hemorrhages and Shaken Baby Syndrome: An Evidence-Based Review. *J. Emerg. Med.* **2008**, *37*, 98–106. [[CrossRef](#)]
12. Falavarjani, K.G.; Parvaresh, M.M.; Shahraki, K.; Nekoozadeh, S.; Amirfarhangi, A. Central retinal artery occlusion in Crohn disease. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* **2012**, *16*, 392–393. [[CrossRef](#)] [[PubMed](#)]
13. Binenbaum, G.; Chen, W.; Huang, J.; Ying, G.-S.; Forbes, B.J. The natural history of retinal hemorrhage in pediatric head trauma. *J. Am. Assoc. Pediatr. Ophthalmol. Strabismus* **2016**, *20*, 131–135. [[CrossRef](#)]
14. Sengupta, S.; Singh, A.; Leopold, H.A.; Gulati, T.; Lakshminarayanan, V. Ophthalmic diagnosis using deep learning with fundus images—A critical review. *Artif. Intell. Med.* **2019**, *102*, 101758. [[CrossRef](#)]
15. Balyen, L.; Peto, T. Promising Artificial Intelligence-Machine Learning-Deep Learning Algorithms in Ophthalmology. *Asia-Pacific J. Ophthalmol.* **2019**, *8*, 264–272. [[CrossRef](#)]

16. Ting, D.S.; Peng, L.; Varadarajan, A.V.; Keane, P.A.; Burlina, P.M.; Chiang, M.F.; Schmetterer, L.; Pasquale, L.R.; Bressler, N.M.; Webster, D.R.; et al. Deep learning in ophthalmology: The technical and clinical considerations. *Prog. Retin. Eye Res.* **2019**, *72*, 100759. [[CrossRef](#)] [[PubMed](#)]
17. Ting, D.S.W.; Pasquale, L.R.; Peng, L.; Campbell, J.P.; Lee, A.Y.; Raman, R.; Tan, G.S.W.; Schmetterer, L.; Keane, P.A.; Wong, T.Y. Artificial intelligence and deep learning in ophthalmology. *Br. J. Ophthalmol.* **2019**, *103*, 167–175. [[CrossRef](#)]
18. Tsiknakis, N.; Theodoropoulos, D.; Manikis, G.; Ktistakis, E.; Boutsora, O.; Berto, A.; Scarpa, F.; Scarpa, A.; Fotiadis, D.I.; Marias, K. Deep learning for diabetic retinopathy detection and classification based on fundus images: A review. *Comput. Biol. Med.* **2021**, *135*, 104599. [[CrossRef](#)] [[PubMed](#)]
19. Wu, J.-H.; Liu, T.Y.A. Application of Deep Learning to Retinal-Image-Based Oculomics for Evaluation of Systemic Health: A Review. *J. Clin. Med.* **2022**, *12*, 152. [[CrossRef](#)]
20. Panda, N.R.; Sahoo, A.K. A Detailed Systematic Review on Retinal Image Segmentation Methods. *J. Digit. Imaging* **2022**, *35*, 1250–1270. [[CrossRef](#)]
21. Moradi, M.; Chen, Y.; Du, X.; Seddon, J.M. Deep ensemble learning for automated non-advanced AMD classification using optimized retinal layer segmentation and SD-OCT scans. *Comput. Biol. Med.* **2023**, *154*, 106512. [[CrossRef](#)] [[PubMed](#)]
22. Prahs, P.; Radeck, V.; Mayer, C.; Cvetkov, Y.; Cvetkova, N.; Helbig, H.; Märker, D. OCT-based deep learning algorithm for the evaluation of treatment indication with anti-vascular endothelial growth factor medications. *Graefes Arch. Clin. Exp. Ophthalmol.* **2017**, *256*, 91–98. [[CrossRef](#)]
23. Campbell, J.P.; Chiang, M.F.; Chen, J.S.; Moshfeghi, D.M.; Nudleman, E.; Cherwek, H.; Cheung, C.Y.; Singh, P.; Kalpathy-Cramer, J.; Ostmo, S.; et al. Artificial Intelligence for Retinopathy of Prematurity. *Ophthalmology* **2022**, *129*, e69–e76. [[CrossRef](#)]
24. Campbell, J.P.; Singh, P.; Redd, T.K.; Brown, J.M.; Shah, P.K.; Subramanian, P.; Rajan, R.; Valikodath, N.; Cole, E.; Ostmo, S.; et al. Applications of Artificial Intelligence for Retinopathy of Prematurity Screening. *Pediatrics* **2021**, *147*, e2020016618. [[CrossRef](#)] [[PubMed](#)]
25. Faust, O.; Acharya, U.R.; Ng, E.Y.K.; Ng, K.H.; Suri, J.S. Algorithms for the Automated Detection of Diabetic Retinopathy Using Digital Fundus Images: A Review. *J. Med. Syst.* **2010**, *36*, 145–157. [[CrossRef](#)] [[PubMed](#)]
26. Huang, X.; Wang, H.; She, C.; Feng, J.; Liu, X.; Hu, X.; Chen, L.; Tao, Y. Artificial intelligence promotes the diagnosis and screening of diabetic retinopathy. *Front. Endocrinol.* **2022**, *13*, 946915. [[CrossRef](#)]
27. Keenan, T.D.; Dharssi, S.; Peng, Y.; Chen, Q.; Agrón, E.; Wong, W.T.; Lu, Z.; Chew, E.Y. A Deep Learning Approach for Automated Detection of Geographic Atrophy from Color Fundus Photographs. *Ophthalmology* **2019**, *126*, 1533–1540. [[CrossRef](#)]
28. Yan, Q.; Weeks, D.E.; Xin, H.; Swaroop, A.; Chew, E.Y.; Huang, H.; Ding, Y.; Chen, W. Deep-learning-based prediction of late age-related macular degeneration progression. *Nat. Mach. Intell.* **2020**, *2*, 141–150. [[CrossRef](#)]
29. Lee, J.; Wanyan, T.; Chen, Q.; Keenan, T.D.L.; Glicksberg, B.S.; Chew, E.Y.; Lu, Z.; Wang, F.; Peng, Y. Predicting Age-related Macular Degeneration Progression with Longitudinal Fundus Images Using Deep Learning. *Mach. Learn. Med. Imaging* **2022**, *13583*, 11–20. [[CrossRef](#)] [[PubMed](#)]
30. Morano, J.; Hervella, S.; Rouco, J.; Novo, J.; Fernández-Vigo, J.I.; Ortega, M. Weakly-supervised detection of AMD-related lesions in color fundus images using explainable deep learning. *Comput. Methods Programs Biomed.* **2023**, *229*, 107296. [[CrossRef](#)] [[PubMed](#)]
31. Li, F.; Su, Y.; Lin, F.; Li, Z.; Song, Y.; Nie, S.; Xu, J.; Chen, L.; Chen, S.; Li, H.; et al. A deep-learning system predicts glaucoma incidence and progression using retinal photographs. *J. Clin. Investig.* **2022**, *132*, e157968. [[CrossRef](#)]
32. Coan, L.J.; Williams, B.M.; Adithya, V.K.; Upadhyaya, S.; Alkafri, A.; Czanner, S.; Venkatesh, R.; Willoughby, C.E.; Kavitha, S.; Czanner, G. Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review. *Surv. Ophthalmol.* **2022**, *68*, 17–41. [[CrossRef](#)] [[PubMed](#)]
33. Zhang, L.; Tang, L.; Xia, M.; Cao, G. The application of artificial intelligence in glaucoma diagnosis and prediction. *Front. Cell Dev. Biol.* **2023**, *11*, 1173094. [[CrossRef](#)] [[PubMed](#)]
34. Gutierrez, L.; Lim, J.S.; Foo, L.L.; Ng, W.Y.Y.; Yip, M.; Lim, G.Y.S.; Wong, M.H.Y.; Fong, A.; Rosman, M.; Mehta, J.S.; et al. Application of artificial intelligence in cataract management: Current and future directions. *Eye Vis.* **2022**, *9*, 1–11. [[CrossRef](#)]
35. Park, Y.-J.; Bae, J.H.; Shin, M.H.; Hyun, S.H.; Cho, Y.S.; Choe, Y.S.; Choi, J.Y.; Lee, K.-H.; Kim, B.-T.; Moon, S.H. Development of Predictive Models in Patients with Epiphora Using Lacrimal Scintigraphy and Machine Learning. *Nucl. Med. Mol. Imaging* **2019**, *53*, 125–135. [[CrossRef](#)] [[PubMed](#)]
36. Kuo, B.-I.; Chang, W.-Y.; Liao, T.-S.; Liu, F.-Y.; Liu, H.-Y.; Chu, H.-S.; Chen, W.-L.; Hu, F.-R.; Yen, J.-Y.; Wang, I.-J. Keratoconus Screening Based on Deep Learning Approach of Corneal Topography. *Transl. Vis. Sci. Technol.* **2020**, *9*, 53. [[CrossRef](#)] [[PubMed](#)]
37. Chun, J.; Kim, Y.; Shin, K.Y.; Han, S.H.; Oh, S.Y.; Chung, T.-Y.; Park, K.-A.; Lim, D.H. Deep Learning-Based Prediction of Refractive Error Using Photorefractive Images Captured by a Smartphone: Model Development and Validation Study. *JMIR Public Heal. Surveill.* **2020**, *8*, e16225. [[CrossRef](#)]
38. Dong, L.; He, W.; Zhang, R.; Ge, Z.; Wang, Y.X.; Zhou, J.; Xu, J.; Shao, L.; Wang, Q.; Yan, Y.; et al. Artificial Intelligence for Screening of Multiple Retinal and Optic Nerve Diseases. *JAMA Netw. Open* **2022**, *5*, e229960. [[CrossRef](#)] [[PubMed](#)]
39. Milea, D.; Najjar, R.P.; Jiang, Z.; Ting, D.; Vasseneix, C.; Xu, X.; Fard, M.A.; Fonseca, P.; Vanikieti, K.; Lagrèze, W.A.; et al. Artificial Intelligence to Detect Papilledema from Ocular Fundus Photographs. *N. Engl. J. Med.* **2020**, *382*, 1687–1695. [[CrossRef](#)] [[PubMed](#)]
40. Aziz, T.; Charoenlarnpparat, C.; Mahapakulchai, S. Deep learning-based hemorrhage detection for diabetic retinopathy screening. *Sci. Rep.* **2023**, *13*, 1479. [[CrossRef](#)] [[PubMed](#)]

41. Skouta, A.; Elmoufidi, A.; Jai-Andaloussi, S.; Ouchetto, O. Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network. *J. Big Data* **2022**, *9*, 78. [[CrossRef](#)]
42. Mao, J.; Luo, Y.; Chen, K.; Lao, J.; Chen, L.; Shao, Y.; Zhang, C.; Sun, M.; Shen, L. New grading criterion for retinal haemorrhages in term newborns based on deep convolutional neural networks. *Clin. Exp. Ophthalmol.* **2020**, *48*, 220–229. [[CrossRef](#)]
43. Galdran, A.; Anjos, A.; Dolz, J.; Chakor, H.; Lombaert, H.; Ayed, I.B. State-of-the-art retinal vessel segmentation with minimalistic models. *Sci. Rep.* **2022**, *12*, 6174. [[CrossRef](#)]
44. Balasubramanian, V. Statistical Inference, Occam’s Razor, and Statistical Mechanics on the Space of Probability Distributions. *Neural Comput.* **1997**, *9*, 349–368. [[CrossRef](#)]
45. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255. [[CrossRef](#)]
46. Devries, T.; Taylor, G.W. Improved Regularization of Convolutional Neural Networks with Cutout. *arXiv* **2017**, arXiv:1708.04552.
47. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [[CrossRef](#)]
48. Wang, F.; Jiang, M.; Qian, C.; Yang, S.; Li, C.; Zhang, H.; Wang, X.; Tang, X. Residual attention network for image classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3156–3164.
49. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16 × 16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.
50. Vasu, P.K.A.; Gabriel, J.; Zhu, J.; Tuzel, O.; Ranjan, A. FastViT: A Fast Hybrid Vision Transformer using Structural Reparameterization. *arXiv* **2023**, arXiv:2303.14189.
51. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Proceedings of the NeurIPS, Vancouver, BC, Canada, 8–14 December 2019; pp. 8024–8035.
52. PyTorch Image Models, 2019, 10.5281/zenodo. 4414861. GitHub Repository. Available online: <https://github.com/rwightman/pytorch-image-models> (accessed on 1 September 2023).
53. Steiner, A.; Kolesnikov, A.; Zhai, X.; Wightman, R.; Uszkoreit, J.; Beyer, L. How to train your vit? data, augmentation, and regularization in vision transformers. *arXiv* **2021**, arXiv:2106.10270.
54. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
55. Bergstra, J.; Yamins, D.; Cox, D.D. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. In Proceedings of the 30th International Conference on Machine Learning (ICML 2013), Atlanta, GA, USA, 16–21 June 2013; pp. I-115–I-123.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.