Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Facial Photograph–Based Detection of Dermatochalasis and Assessment of Blepharoplasty Necessity: A Deep Learning–Assisted Approach

Hakan Veli SAVAŞ

hakanvelisavas@hotmail.com

Elazığ Medical Hospital

Osman ALTAY

Manisa Celal Bayar University

Research Article

Keywords: Dermatochalasis, Blepharoplasty, Deep learning, Convolutional neural network, Facial photographs

Posted Date: August 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9341023/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

To develop and compare convolutional neural network (CNN) models for distinguishing upper eyelid blepharoplasty candidates from controls using routine frontal facial photographs.

Methods

In this single-center diagnostic modelling study, 380 adults were included: 190 blepharoplasty candidates and 190 healthy controls with normal-appearing upper eyelids. Standardized frontal facial photographs were obtained under outpatient conditions using a smartphone camera. Three ImageNet-pretrained CNN architectures (ResNet18, EfficientNet-B0, and DenseNet121) were fine-tuned for binary classification under three augmentation regimes. Performance was evaluated on an independent test set using accuracy, recall, specificity, F1-score, Cohen's κ, area under the receiver operating characteristic curve (AUC), and total runtime.

Results

Across all architectures and augmentation regimes, the models showed very high discriminatory performance, with overall accuracies of approximately 98% and recall for the blepharoplasty candidate class reaching 100%. AUC values ranged from 0.9791 to 1.0000. EfficientNet-B0 under the no_noise regime achieved the highest AUC (0.9901), whereas ResNet18 required the shortest runtime. EfficientNet-B0 and DenseNet121 were more computationally demanding.

Conclusion

Deep CNNs can accurately distinguish upper eyelid blepharoplasty candidates from controls using only standardized frontal facial photographs. These findings support image-based deep learning as a promising adjunct for more objective and efficient triage in upper eyelid blepharoplasty, although external validation in larger and more heterogeneous populations is required.

Introduction

Dermatochalasis is characterized by excess upper eyelid skin, age-related periocular tissue laxity, degeneration of collagen and elastin fibers, reduced dermal hydration, and progressive weakening of the periocular soft tissues.[1, 2] Its pathophysiology involves thinning of the dermal connective tissue and attenuation of the orbital septum, anterior prolapse of the preaponeurotic fat pads, chronic ultraviolet-induced photoaging, decreased hyaluronic acid content leading to volume and moisture loss, and reduced synthesis of type I and type III collagen accompanied by fragmentation of elastic fibers. [3–5]

Clinically, dermatochalasis manifests as upper eyelid heaviness, superior visual field restriction, visual fatigue, a subjective feeling of eyelid drooping, and compensatory frontalis overaction leading to brow and forehead fatigue, together with a characteristic "tired" and aged facial appearance.[6–8] These symptoms can negatively impact both aesthetic satisfaction and functional visual performance in daily activities.

Blepharoplasty is one of the most commonly performed oculoplastic procedures worldwide for both functional and cosmetic indications, and increasing patient demand has made it a substantial contributor to healthcare expenditure.[9] According to the 2023 report from the American Society of Plastic Surgeons (ASPS), more than 115,000 upper eyelid blepharoplasty procedures are performed annually in the United States, with an estimated yearly economic burden of 300–400 million USD.[10, 11] These figures highlight the high clinical prevalence of dermatochalasis and the considerable economic impact of blepharoplasty on modern healthcare systems.

Several studies have reported significant improvements in superior visual field, brow and forehead fatigue, and patient-reported quality-of-life outcomes following upper eyelid blepharoplasty.[9, 12] These findings underscore the importance of accurately identifying suitable surgical candidates. In routine clinical practice, the indication for blepharoplasty is largely determined by the clinician's overall impression, the patient's symptoms, and photographic documentation. Although standard measurements such as margin reflex distance 1 (MRD1), eyelid crease height, and the amount of excess skin fold are commonly employed, these parameters can exhibit considerable interobserver variability and may not fully capture the patient's actual functional burden or degree of aesthetic dissatisfaction.[13, 14] As the demand for upper eyelid rejuvenation continues to rise, the routine acquisition of high-resolution facial photographs in clinical settings, together with the widespread availability of smartphone cameras, provides a unique opportunity for image-based automated analysis systems. In this context, artificial intelligence– and deep learning–based decision-support tools that quantitatively evaluate facial and eyelid photographs have the potential to complement traditional subjective examination, reduce interobserver variability, standardize clinical decision-making, and make patient counselling more consistent and predictable, thereby anchoring blepharoplasty indications not only in clinical intuition but also in algorithmically defined, data-driven, and reproducible criteria.

Over the past decade, deep learning–based image analysis has reshaped ophthalmology by enabling automated detection, classification, and prognostication across a wide spectrum of disorders, including diabetic retinopathy, age-related macular degeneration, glaucoma, and cataract. Convolutional neural networks (CNNs), in particular, have achieved performance levels comparable to expert human graders in large-scale screening and diagnostic tasks and are increasingly being integrated into real-world clinical workflows.[15, 16] Building on these advances in posterior segment imaging, recent studies have extended deep learning applications to anterior segment and periocular photographs, enabling automated recognition of ocular surface disease, pediatric glaucoma, and various eyelid and orbital pathologies using simple two-dimensional images.[17, 18]

Within the oculoplastic domain, several groups have investigated the use of AI for quantitative analysis of eyelid morphology and periocular appearance. Deep learning and CNN-based models have been used to measure eyelid parameters, evaluate outcomes after blepharoptosis surgery, and detect referable ptosis from facial photographs with accuracy comparable to non-ophthalmologist physicians.[19–21] Other AI-driven tools estimate apparent facial age and cosmetic change before and after eyelid procedures, offering objective evidence of rejuvenating effects.[22–24] However, despite this growing body of work, most existing approaches target postoperative outcome assessment or ptosis detection rather than the preoperative question of whether an individual should be considered a candidate for upper eyelid blepharoplasty based solely on standard frontal facial photographs. To our knowledge, no previous study has systematically evaluated commonly used CNN architectures on mixed datasets comprising healthy individuals and real-world blepharoplasty candidates to develop a practical, image-based decision-support model for blepharoplasty candidacy.

In this context, facial photographs represent a highly attractive and widely accessible input modality for the automated assessment of blepharoplasty candidacy. In the present study, we develop and benchmark a deep learning–based decision-support model that performs a binary classification of individuals as potential candidates for upper eyelid blepharoplasty or as having clinically normal upper eyelids, using only standard frontal facial photographs as input. To approximate real-world clinical conditions, our dataset combines images of healthy individuals from an openly available internet-based facial image repository with photographs of patients presenting to an oculoplastic clinic with upper eyelid dermatochalasis and a clinically suspected indication for blepharoplasty; all clinical images were captured under routine outpatient conditions using a standard straight-ahead facial view, mirroring the photographs commonly documented in everyday practice. Within an identical training and evaluation pipeline, we systematically compare three widely used convolutional neural network architectures—ResNet18, EfficientNet-B0, and DenseNet121. In view of a growing global population, the rising demand for aesthetic and functional eyelid surgery, and persistent inequities in access to specialist care, it is essential that such models remain computationally efficient and deployable on resource-limited platforms, including mobile applications and handheld devices. Accordingly, for each architecture we evaluate not only conventional classification metrics but also overall runtime, with the goal of identifying a baseline model that balances diagnostic performance with computational efficiency and is practically deployable in real-world settings. To our knowledge, this is the first study to address the question of blepharoplasty candidacy directly from routine facial photographs by jointly evaluating multiple CNN architectures on a mixed dataset of healthy controls and real-world blepharoplasty candidates within a unified, clinically oriented framework explicitly designed for real-world use and potential mobile implementation.

Materials and Methods

This single-center diagnostic modelling study was conducted in the ophthalmology outpatient clinic of Karakoçan State Hospital, Elazığ Province, Turkey, where all standardized frontal facial photographs were obtained, between April 2025 and August 2025. All clinical procedures were performed in accordance with the tenets of the Declaration of Helsinki, and the study protocol was approved by the Non-Interventional Research Ethics Committee of Elazığ Fethi Sekin City Hospital at its meeting on 16 January 2025 (decision no. 2025/2–6). Written informed consent for clinical facial photography and subsequent anonymized image analysis was obtained from all participating patients prior to enrolment.

The blepharoplasty candidate group comprised adult patients who presented to the oculoplastic outpatient clinic with upper eyelid dermatochalasis and were clinically judged to be potential candidates for upper eyelid blepharoplasty. Patients were eligible if they had sufficient excess upper eyelid skin to be considered for surgery by the attending ophthalmologist and if adequate standard frontal facial photographs could be obtained. Exclusion criteria included a history of prior periocular or facial surgery affecting the upper eyelids or brows, evident neurogenic or myogenic ptosis, marked facial asymmetry, significant facial nerve dysfunction, or any condition that prevented reliable photographic documentation (e.g., inability to maintain primary gaze or keep the eyes open during image capture).

To obtain a comparison group with clinically normal upper eyelids, healthy control images were selected from a publicly available stock photo repository (Shutterstock.com; Shutterstock, Inc., New York, NY, USA). Only adults without obvious upper eyelid dermatochalasis, ptosis, significant brow ptosis, or other visible periocular pathology were included based on visual inspection. Images were excluded if they showed extreme facial expressions; heavy makeup obscuring the upper eyelid region; occlusions such as glasses, hair, or accessories covering the eyelids; or non-frontal facial poses.

All clinical facial photographs were obtained in the same oculoplastic outpatient clinic using a smartphone-based digital imaging system (iPhone 16 Pro Max, Apple Inc., Cupertino, CA, USA). Images were captured in a standard straight-ahead facial view with the patient seated at eye level, looking directly into the camera in primary gaze, with a neutral facial expression and both eyes fully open. The camera-to-patient distance and framing were adjusted to include the full face while ensuring that the periocular region was clearly visible. Ambient illumination reflected routine outpatient conditions, and no additional external lighting equipment was used beyond the device's integrated system. Before being used for model development, all images were anonymized and assigned random identification codes. Any metadata or facial information that could reveal patient identity was removed.

Data Augmentation

To enhance robustness to variations in illumination, pose, blur, and sensor-induced noise, three controlled data-augmentation regimes were applied during training.[25] All images were resized to 224 × 224 pixels, normalized using a mean of 0.5 and a standard deviation of 0.5, and split into a stratified 70/15/15 train/validation/test set. Images in the validation and test sets were not augmented. Data augmentation was applied exclusively to the training set using the following three modes:

- **full_aug**: Combined geometric transformations (random horizontal flip, ± 10° rotation) and photometric adjustments (brightness and contrast jitter), together with Gaussian blur and mild additive Gaussian noise.[26]
- **no_blur**: Identical to *full_aug* but without blur, enabling isolated assessment of blur-related regularization.[27]
- **no_noise**: Identical to *full_aug* but without additive noise, allowing independent evaluation of noise-based regularization.[28]

These three augmentation regimes were applied uniformly across all CNN architectures (ResNet18, EfficientNet-B0, DenseNet121), ensuring that the impact of blur and noise could be compared under identical preprocessing and training conditions.

CNN Architectures

Three widely used convolutional neural network (CNN) architectures—ResNet18, EfficientNet-B0, and DenseNet121—were evaluated as baseline models. All networks were initialized with ImageNet-pretrained weights and adapted for binary classification by replacing the final fully connected layer with a two-unit output layer; no additional architectural modifications were introduced.

Despite their structural differences—ResNet making use of residual skip connections,[29] EfficientNet employing MBConv and squeeze-and-excitation blocks,[30] and DenseNet relying on densely connected feature propagation [31] —the training pipeline was kept identical for all models. Each network received the same input resolution (224 × 224 pixels), normalization, data split, and data-augmentation regime, differing only in backbone architecture.

All models were trained end-to-end with the Adam optimizer (learning rate 1×10^{-4}), a batch size of 32, and cross-entropy loss, with full fine-tuning of all layers for 5 epochs.[32] This standardized configuration was chosen to ensure that any observed performance differences reflected the intrinsic capabilities of each CNN architecture rather than differences in preprocessing or hyperparameter settings.

Results

Demographic and Clinical Characteristics

The dataset included 380 adults: 190 blepharoplasty candidates with clinically evident upper eyelid dermatochalasis and 190 controls with normal-appearing upper eyelids. Candidates had a mean age of 58.1 ± 8.9 years and were 68.4% female, while controls had a mean age of 52.6 ± 9.1 years and were 69.4% female, with no clinically relevant differences in age or sex between groups. All facial photographs were obtained in a single oculoplastic outpatient clinic using a smartphone camera at eye level, with patients in primary gaze, neutral expression, and both eyes fully open under ambient room lighting. These routine, non-studio conditions were chosen to mirror everyday clinical workflow and support the external validity of the models.

This section presents a comprehensive evaluation of ResNet18, EfficientNet-B0, and DenseNet121 under three data-augmentation regimes (full_aug, no_blur, and no_noise), integrating quantitative performance metrics, learning dynamics, receiver operating characteristic (ROC) behavior, and computational efficiency. All metrics are reported on the independent test set.

As summarized in Table 1, all three architectures achieve nearly identical classification performance across all augmentation conditions. Accuracy (0.9825), precision (0.9655), recall (1.0000), F1-score (0.9825), specificity (0.9655), and Cohen's κ (0.9649) remain constant for every model–augmentation combination. These findings indicate that, in our study, the classification task is highly separable and that differences in architectural depth or connectivity do not translate into measurable gains in predictive performance.

Table 1
Overall classification performance and inference time for ResNet18, EfficientNet-B0, and DenseNet121 across the three data-augmentation regimes (full_aug, no_blur, no_noise) on the independent test set.

Mode	Model	Accuracy	Precision	Recall	F1	Specificity	AUC	Kappa	Duration_sec
	resnet18	**0.9825**	**0.9655**	**1.0000**	**0.9825**	**0.9655**	**1.0000**	**0.9649**	**171.98**
full_aug	efficientnet_b0	0.9825	0.9655	1.0000	0.9825	0.9655	0.9791	0.9649	206.64
	densenet121	0.9825	0.9655	1.0000	0.9825	0.9655	1.0000	0.9649	289.05
	resnet18	0.9825	0.9655	1.0000	0.9825	0.9655	1.0000	0.9649	169
no_blur	efficientnet_b0	0.9825	0.9655	1.0000	0.9825	0.9655	0.9815	0.9649	205.65
	densenet121	0.9825	0.9655	1.0000	0.9825	0.9655	1.0000	0.9649	287.87
	resnet18	0.9825	0.9655	1.0000	0.9825	0.9655	0.9963	0.9649	171.05
no_noise	efficientnet_b0	0.9825	0.9655	1.0000	0.9825	0.9655	0.9901	0.9649	206.28
	densenet121	0.9825	0.9655	1.0000	0.9825	0.9655	0.9975	0.9649	286.31

Although threshold-based metrics do not differentiate between the models, small yet consistent variations are observed in the area under the ROC curve (AUC). As summarized in Table 1 and illustrated in Fig. 3, AUC values remain very high (0.9791–1.0000) for all models. EfficientNet-B0 shows a slight improvement under the *no_noise* configuration (AUC = 0.9901), whereas ResNet18 and DenseNet121 remain close to 1.000 across all augmentation regimes. These fluctuations correspond to only marginal differences in ROC curvature and do not affect the final classification outcomes, confirming that all architectures operate near the optimal region of the ROC space.

Learning dynamics further support these observations. Figure 1 depicts the training and validation curves for each augmentation setup, demonstrating rapid and stable convergence in all architectures without evidence of overfitting. The similarity of accuracy and loss trajectories across augmentation types indicates that blur, noise, and spatial transformations do not hinder the extraction of discriminative eyelid features. This behavior is consistent with the uniformity observed in Table 1 and suggests that the dataset supports robust feature learning regardless of the chosen augmentation regime.

Beyond classification metrics, computational efficiency is the only dimension in which the architectures differ meaningfully. As reported in Table 1 and visualized in Fig. 2, ResNet18 achieves the shortest runtime (169–172 seconds), EfficientNet-B0 demonstrates intermediate runtimes (205–207 seconds), and DenseNet121 requires the longest processing time (286–289 seconds). Given that predictive performance remains indistinguishable across models, runtime becomes the principal factor guiding model selection. These empirical observations indicate that, for this task, increased architectural complexity does not yield measurable gains in predictive performance but instead incurs additional computational cost.

Taken together, the evidence from Table 1 and Figs. 1–3 demonstrates that the three architectures are essentially equivalent in diagnostic performance but differ substantially in computational requirements. Consequently, computational efficiency, rather than accuracy, becomes the key determinant in practical deployment scenarios.

The observed runtime profiles suggest different levels of suitability depending on hardware capacity and operational constraints. The markedly shorter runtime of ResNet18 makes it a strong candidate for environments in which inference speed or resource limitations are critical, such as mobile-based screening tools, portable diagnostic units, or embedded systems. EfficientNet-B0, with its moderate computational demand, represents a balanced option for settings that rely on server-assisted workflows or hybrid mobile–cloud processing, where latency is important but not strictly real time. DenseNet121, which carries the highest computational load without offering additional diagnostic benefit, is most appropriate for offline batch analysis or high-resource environments in which runtime constraints are less restrictive. Overall, the results indicate that all three models provide reliable diagnostic accuracy, and that the optimal choice depends primarily on deployment constraints rather than on differences in predictive performance.

From a clinical perspective, the proposed model exhibits a performance profile that is highly reassuring for the screening and triage of patients with upper eyelid dermatochalasis. In our cohort, it correctly identified all patients who had been judged by oculoplastic surgeons as blepharoplasty candidates, while only occasionally flagging healthy individuals as possible candidates. This pattern—prioritizing the avoidance of missed clinically relevant dermatochalasis cases at the expense of a small number of additional referrals—closely mirrors real-world oculoplastic priorities, in which overlooking a functionally or cosmetically significant case is considered more problematic than referring a borderline patient for further face-to-face evaluation.

Discussion

Our study aims to develop an objective deep learning–based framework to identify candidates for upper eyelid blepharoplasty by detecting clinically significant dermatochalasis. Dermatochalasis, in which redundant upper eyelid skin impairs appearance and visual function, is both common and costly; in the United States alone, more than 110,000 upper eyelid procedures are performed each year, with an estimated annual cost of about 400 million USD [10, 11] Yet the decision to offer surgery remains largely subjective, relying on clinician judgment and variably measured indices such as margin reflex distance and skin-fold height. Artificial intelligence (AI) is already widely used in ophthalmology—especially in retinal disease, with more than 3,000 publications on “retina and AI”—but oculoplastic eyelid disorders remain relatively underexplored: as of December 2025, PubMed lists only 49 studies combining the terms “oculoplastics” and “AI.” Against this backdrop, our work, to our knowledge, is the first to specifically develop and rigorously evaluate an AI-based tool to support blepharoplasty candidacy decision-making using routine frontal facial photographs.

In this study, we show that deep convolutional neural networks can accurately distinguish upper eyelid blepharoplasty candidates from individuals with normal eyelids using only standard frontal facial photographs. All three architectures (ResNet18, EfficientNet-B0, and DenseNet121) achieved similarly high performance across augmentation strategies, with accuracies around 98% and 100% recall for the candidate class on the independent test set, indicating a strong and consistently learnable morphological signal rather than an architecture-specific artifact. Under these conditions, computational efficiency becomes the main criterion for model choice: ResNet18 required substantially less runtime and is attractive for mobile or other resource-limited deployments, whereas the heavier backbones are better suited to server-based or batch-processing applications.

Our findings are consistent with and extend a growing body of work demonstrating that AI can reliably quantify eyelid morphology, detect ptosis, and evaluate outcomes from facial or periocular images. Lou et al.[33], Nam et al.[34], Shao et al.[35], and Rana et al.[36] have reported deep learning–based systems that automatically measure key eyelid parameters with near-manual accuracy and extract periocular anthropometric landmarks from routine clinical photographs, thereby supporting objective, image-based assessment of eyelid position and contour. In addition, several groups have developed ptosis-screening models—such as the smartphone video–based MRD1 estimator proposed by Lootus et al.[37] —that can identify clinically significant ptosis with performance comparable to that of non-specialist physicians. In parallel, AI-driven tools have been applied to cosmetic and functional outcome evaluation after eyelid surgery. These include algorithms that estimate apparent age and attractiveness and demonstrate a modest but measurable rejuvenation effect following upper eyelid procedures, as well as models that predict postoperative appearance in blepharoptosis surgery and support preoperative counseling and expectation management.[38–42] Taken together, these studies indicate that eyelid morphology carries a strong, quantifiable signal that can be captured by AI-based methods. Our work builds on this foundation by moving beyond measurement and diagnosis toward a more integrative endpoint—namely, determining whether a given patient should be considered a candidate for upper eyelid blepharoplasty.

Recent reviews of artificial intelligence in oculoplastics and ophthalmic surgery have highlighted that AI systems can now diagnose, quantify, and predict outcomes across a wide range of eyelid and orbital conditions, yet candidate-screening tools remain underdeveloped.[43–46] By focusing on the upstream task of objectively triaging and prioritizing patients who are likely to benefit from upper eyelid blepharoplasty, our model suite is well positioned to complement these existing technologies and to streamline the overall clinical decision-making pathway.

Several prior studies have shown that high performance on eyelid and anterior segment AI tasks is achievable, but often with greater variability across model architectures and experimental settings than we observed in our study. Hung et al.[47] reported

that a non–pre-trained DenseNet121 achieved 90.1% sensitivity and 82.4% specificity for blepharoptosis detection, whereas a pre-trained ResNet34 performed markedly worse, with sensitivity and specificity falling to 74.1% and 63.6%, and mean accuracy remaining in the low-to-mid 80% range. Likewise, Tabuchi et al.[48] obtained an AUC of 0.90 for tablet-based ptosis screening, yet sensitivity and specificity hovered around 83% with noticeable variability across folds. In contrast, in our classification system, the performance curves of all three CNN architectures under all augmentation strategies were nearly superimposable. This high degree of consistency indicates that, regardless of the specific architecture used, our models tend to make the same blepharoplasty candidacy decisions for the same patients, underscoring the robustness and practical generalizability of our approach.

Other oculoplastic and anterior segment AI studies show that performance often drops when tasks are less constrained or image quality and case mix are more variable. Chen et al.[49] developed a smartphone system to estimate MRD1, MRD2 and levator function, with only moderate agreement for some measures and a focus on metric regression rather than surgical decisions. Kuo et al.[50] tested several CNNs for bacterial keratitis from external photographs and, despite optimization, achieved only modest accuracy without clear benefit of EfficientNet-B3. A recent systematic review of AI using external facial or ocular images likewise reported wide accuracy ranges and strong dependence on label quality, population characteristics and imaging conditions.[51] Together, these data suggest that our model benefits from a tightly defined task—standardized frontal photographs, a single clinically meaningful phenotype and a binary “operate vs do not operate” label based on expert consensus—while also implying that our results represent an optimistic upper bound that will need rigorous external and prospective validation before clinical deployment.[51]

Beyond numerical performance, the models’ behavior is also clinically reassuring: across analyses they consistently distinguished patients with clinically meaningful dermatochalasis from individuals with normal eyelids, suggesting that the key periocular features guiding blepharoplasty decisions are reliably captured. In practice, this tool could serve as an initial screening aid in outpatient or telemedicine settings—highlighting patients who may merit oculoplastic referral, supporting standardized photographic triage, and reducing interobserver variability in referral thresholds, while complementing rather than replacing clinical examination.

This study has several limitations. First, it used single-center data from standardized frontal facial photographs acquired under relatively controlled conditions, which may limit generalizability to more diverse populations, camera systems, lighting conditions, and telemedicine settings. Second, although age and sex distributions were comparable, blepharoplasty candidates and controls came from different sources (clinical photographs vs stock images), so subtle differences in image quality or facial presentation may have influenced the models in addition to dermatochalasis-related features. Third, ground-truth labels were based on expert clinical judgment rather than functional outcomes or patient-reported measures, meaning the networks largely reproduce existing decision patterns, including potential biases. Finally, we did not perform external or prospective validation, evaluate performance across demographic subgroups, or compare our CNNs with simpler rule-based or measurement-driven models, so the present results should be interpreted as proof of concept rather than evidence of clinical superiority.

Conclusion

This study shows that deep convolutional neural networks can distinguish upper eyelid blepharoplasty candidates from controls with very high accuracy using routine frontal facial photographs. These models may serve as an objective adjunct to standard clinical assessment and support photographic triage, but require validation in larger and more diverse populations before routine use.

Declarations

Competing Interests:

The authors have no relevant financial or non-financial interests to disclose.

Ethics Approval:

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Elazığ Fethi Sekin City Hospital (16 January 2025; Decision No. 2025/2–6).

Consent to Participate:

Written informed consent was obtained from all individual participants included in the study.

Consent to Publish:

The authors affirm that all participants provided informed consent for publication of any images or other potentially identifiable material included in this article.

Funding:

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Author Contribution

Author Contributions: All authors contributed to the study conception and design. Data collection, data analysis, and interpretation were performed by Hakan Veli Savaş and Osman Altay. The first draft of the manuscript was written by Hakan Veli Savaş, and both authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. Hakan Veli Savaş and Osman Altay acted as guarantors of the work.

Data Availability:

The datasets generated and/or analysed during the current study are not publicly available because they include facial photographs and associated clinical data that could compromise participant privacy. De-identified data may be available from the corresponding author on reasonable request, subject to approval by the institutional ethics committee and applicable data protection requirements.

References

1. Verner I, Naveh HP, Cotofana S. A novel ablative radiofrequency microplasma nonsurgical blepharoplasty for dermatochalasis. Dermatol Ther. 2020;33(6). https://doi.org/10.1111/DTH.14002
2. Hsu JTS. Non-Surgical Correction of Dermatochalasis Using Microplasma. J Drugs Dermatol. 2025;24(2):134–138. https://doi.org/10.36849/JDD.8635
3. Shin JW, Kwon SH, Choi JY, et al. Molecular mechanisms of dermal aging and antiaging approaches. Int J Mol Sci. 2019;20(9). https://doi.org/10.3390/IJMS20092126
4. McCabe MC, Hill RC, Calderone K, et al. Alterations in extracellular matrix composition during aging and photoaging of the skin. Matrix Biol Plus. 2020;8. https://doi.org/10.1016/j.mbplus.2020.100041
5. Quan T, Fisher GJ. Role of age-associated alterations of the dermal extracellular matrix microenvironment in human skin aging: A mini-review. Gerontology. 2015;61(5):427–434. https://doi.org/10.1159/000371708
6. Kim JB, Kim JC, Jin HB, Kim SO, Chung JH. Sub-brow blepharoplasty for simultaneous correction of dermatochalasis with upper eyelid hooding and sunken eyelids. Arch Aesthetic Plast Surg. 2019;25(1):27–31. https://doi.org/10.14730/AAPS.2019.25.1.27

7. Ahmadikhatir S, Finskas O, Svedberg K. Evaluation of Headache-Related Quality of Life and Dry Eyes Following Dermatochalasis Surgery: A Prospective Clinical Study. Clin Ophthalmol. 2025;19:2513–2519. https://doi.org/10.2147/OPTH.S520052
8. Maione L, Vinci V, Costanzo D, Battistini A, Lisa A, Di Maria A. Upper eyelid blepharoplasty following hyaluronic acid injection with improved facial aesthetics and eye symptoms: a case report. J Med Case Rep. 2021;15(1). https://doi.org/10.1186/S13256-020-02641-0
9. Hollander MHJ, Contini M, Pott JW, Vissink A, Schepers RH, Jansma J. Functional outcomes of upper eyelid blepharoplasty: A systematic review. J Plast Reconstr Aesthet Surg. 2019;72(2):294–309. https://doi.org/10.1016/j.bjps.2018.11.010
10. Noemí D, Gazcón P, Diamar J, et al. Upper Eyelid Blepharoplasty: An Aesthetic and Functional Repair. Int J Med Sci Clin Res Stud. 2024;4(06):1188–1189. https://doi.org/10.47191/IJMSCRS/V4-I06-30
11. Lee TC, Fung SE, Hu JQ, et al. Is Blepharoplasty Cost-effective? Utility Analysis of Dermatochalasis and Cost-effectiveness Analysis of Upper Eyelid Blepharoplasty. Ophthalmic Plast Reconstr Surg. 2024;40(5):552–559. https://doi.org/10.1097/IOP.0000000000002649
12. Vasović DD, Karamarković ML, Jovanović M, et al. Comprehensive Evaluation of Quality of Life following Upper Eyelid Blepharoplasty: A Prospective Analysis. Medicina (Kaunas). 2024;60(3):500. https://doi.org/10.3390/MEDICINA60030500
13. Chen HC, Tzeng SS, Hsiao YC, Chen RF, Hung EC, Lee OK. Smartphone-based artificial intelligence-assisted prediction for eyelid measurements: Algorithm development and observational validation study. JMIR Mhealth Uhealth. 2021;9(10). https://doi.org/10.2196/32444
14. Lou L, Cao J, Wang Y, et al. Deep learning-based image analysis for automated measurement of eyelid morphology before and after blepharoptosis surgery. Ann Med. 2021;53(1):2278–2285. https://doi.org/10.1080/07853890.2021.2009127
15. Grewal PS, Oloumi F, Rubin U, Tennant MTS. Deep learning in ophthalmology: a review. Can J Ophthalmol. 2018;53(4):309–313. https://doi.org/10.1016/j.jcjo.2018.04.019
16. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019;103(2):167–175. https://doi.org/10.1136/BJOPHTHALMOL-2018-313173
17. Kitaguchi Y, Hayakawa R, Kawashima R, et al. Deep-learning approach to detect childhood glaucoma based on periocular photograph. Sci Rep. 2023;13(1):10141. https://doi.org/10.1038/s41598-023-37389-2
18. Bao XL, Sun YJ, Zhan X, Li GY. Orbital and eyelid diseases: The next breakthrough in artificial intelligence? Front Cell Dev Biol. 2022;10:1069248. https://doi.org/10.3389/FCELL.2022.1069248
19. Abascal Azanza C, Barrio-Barrio J, Ramos Cejudo J, Ybarra Arróspide B, Devoto MH. Development and validation of a convolutional neural network to identify blepharoptosis. Sci Rep. 2023;13(1):17585. https://doi.org/10.1038/s41598-023-44686-3
20. Lou L, Cao J, Wang Y, et al. Deep learning-based image analysis for automated measurement of eyelid morphology before and after blepharoptosis surgery. Ann Med. 2021;53(1):2278. https://doi.org/10.1080/07853890.2021.2009127
21. Hung JY, Chen KW, Perera C, et al. An Outperforming Artificial Intelligence Model to Identify Referable Blepharoptosis for General Practitioners. J Pers Med. 2022;12(2). https://doi.org/10.3390/JPM12020283
22. Chiou TW, Yen CI, Hsiao YC, Chen HC. AI Prediction for Post-Lower Blepharoplasty Age Reduction. Aesthet Surg J. 2024;44(12):NP922-NP930. https://doi.org/10.1093/ASJ/SJAE182
23. Stephanian B, Karki S, Debnath K, et al. Role of Artificial Intelligence and Machine Learning in Facial Aesthetic Surgery: A Systematic Review. Facial Plast Surg Aesthet Med. 2024;26(6):679–705. https://doi.org/10.1089/FPSAM.2024.0204
24. Goodyear K, Saffari PS, Esfandiari M, Baugh S, Rootman DB, Karlin JN. Estimating apparent age using artificial intelligence: Quantifying the effect of blepharoplasty. J Plast Reconstr Aesthet Surg. 2023;85:336–343. https://doi.org/10.1016/j.bjps.2023.07.017
25. Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. J Big Data. 2019;6(1):60. https://doi.org/10.1186/S40537-019-0197-0
26. Zhou SK, Greenspan H, Davatzikos C, et al. A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. Proc IEEE. 2021;109(5):820.

https://doi.org/10.1109/JPROC.2021.3054390

27. Hussain Z, Gimenez F, Yi D, Rubin D. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. AMIA Annu Symp Proc. 2018;2017:979. https://pmc.ncbi.nlm.nih.gov/articles/PMC5977656/. Accessed December 5, 2025.
28. Ramírez-Agudelo OH, Gorea N, Reif A, Bonasera L, Karl M. The role of noisy data in improving CNN robustness for image classification. 2025;13606:197–212. https://doi.org/10.1117/12.3063563
29. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proc IEEE Conf Comput Vis Pattern Recognit. 2016;2016:770–778. https://doi.org/10.1109/CVPR.2016.90
30. Tan M, Le QV. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In: 36th International Conference on Machine Learning (ICML 2019). 2019;2019-June:10691–10700. https://arxiv.org/pdf/1905.11946. Accessed December 5, 2025.
31. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. Proc IEEE Conf Comput Vis Pattern Recognit. 2017;2017:2261–2269. https://doi.org/10.1109/CVPR.2017.243
32. Raghu M, Zhang C, Kleinberg J, Bengio S. Transfusion: Understanding Transfer Learning for Medical Imaging. Adv Neural Inf Process Syst. 2019;32. https://arxiv.org/pdf/1902.07208. Accessed December 5, 2025.
33. Abascal Azanza C, Barrio-Barrio J, Ramos Cejudo J, Ybarra Arróspide B, Devoto MH. Development and validation of a convolutional neural network to identify blepharoptosis. Sci Rep. 2023;13(1):17585. https://doi.org/10.1038/s41598-023-44686-3
34. Lou L, Cao J, Wang Y, et al. Deep learning-based image analysis for automated measurement of eyelid morphology before and after blepharoptosis surgery. Ann Med. 2021;53(1):2278. https://doi.org/10.1080/07853890.2021.2009127
35. Nam Y, Song T, Lee J, Lee JK. Development of a neural network-based automated eyelid measurement system. Sci Rep. 2024;14(1):1202. https://doi.org/10.1038/s41598-024-51838-6
36. Shao J, Huang X, Gao T, et al. Deep learning-based image analysis of eyelid morphology in thyroid-associated ophthalmopathy. Quant Imaging Med Surg. 2023;13(3):1592–1604. https://doi.org/10.21037/QIMS-22-551/COIF
37. Rana K, Beecher M, Caltabiano C, et al. Artificial intelligence to automate assessment of ocular and periocular measurements. Eur J Ophthalmol. 2025;35(1):346–351. https://doi.org/10.1177/11206721241249773
38. Lootus M, Beatson L, Atwood L, et al. Development and Assessment of an Artificial Intelligence-Based Tool for Ptosis Measurement in Adult Myasthenia Gravis Patients Using Selfie Video Clips Recorded on Smartphones. Digit Biomark. 2023;7(1):63–73. https://doi.org/10.1159/000531224
39. Goodyear K, Saffari PS, Esfandiari M, Baugh S, Rootman DB, Karlin JN. Estimating apparent age using artificial intelligence: Quantifying the effect of blepharoplasty. J Plast Reconstr Aesthet Surg. 2023;85:336–343. https://doi.org/10.1016/j.bjps.2023.07.017
40. Balel Y. Evaluation of Changes in Facial Attractiveness and Estimated Facial Age After Blepharoplasty with an Artificial Intelligence Algorithm. Eur J Ther. 2023;29(4):883–890. https://doi.org/10.58600/EURJTHER1648
41. Park JY, Yang H, Cho K, et al. The Usefulness of AI-Based Cornea Exposure Rate (CER) Analysis Utilizing the Anigma View System in Evaluating Ptosis Surgery Outcomes. J Clin Med. 2025;14(5):1691. https://doi.org/10.3390/JCM14051691
42. Sun Y, Huang X, Zhang Q, et al. A Fully Automatic Postoperative Appearance Prediction System for Blepharoptosis Surgery with Image-based Deep Learning. Ophthalmol Sci. 2022;2(3). https://doi.org/10.1016/J.XOPS.2022.100169
43. Huang S, Xie J, Yang B, Gao Q, Ye J. PtosisDiffusion: a training-free workflow for precisely predicting post-operative appearance in blepharoptosis patients based on diffusion models. Front Cell Dev Biol. 2024;12:1459336. https://doi.org/10.3389/FCELL.2024.1459336
44. Nuliqiman M, Xu M, Sun Y, et al. Artificial Intelligence in Ophthalmic Surgery: Current Applications and Expectations. Clin Ophthalmol. 2023;17:3499–3511. https://doi.org/10.2147/OPTH.S438127
45. Vongsachang H, Yoon MK. Artificial Intelligence in Oculoplastic Surgery: Current Landscape. Int Ophthalmol Clin. 2025;65(3):35–39. https://doi.org/10.1097/IIO.0000000000000566
46. Ing E, Bondok M. Oculoplastics and Augmented Intelligence: A Literature Review. J Clin Med. 2025;14(19):6875. https://doi.org/10.3390/JCM14196875

47. Meer E, Kao B, Hekmatjah N, Lu J, Winn B, Grob SR. Artificial Intelligence in Oculoplastics: A Review. Ophthalmic Plast Reconstr Surg. 2025;41(4):372–387. https://doi.org/10.1097/IOP.0000000000002868
48. Hung JY, Perera C, Chen KW, et al. A deep learning approach to identify blepharoptosis by convolutional neural networks. Int J Med Inform. 2021;148:104402. https://doi.org/10.1016/J.IJMEDINF.2021.104402
49. Tabuchi H, Nagasato D, Masumoto H, et al. Developing an iOS application that uses machine learning for the automated diagnosis of blepharoptosis. Graefes Arch Clin Exp Ophthalmol. 2022;260(4):1329–1335. https://doi.org/10.1007/S00417-021-05475-8
50. Chen HC, Tzeng SS, Hsiao YC, Chen RF, Hung EC, Lee OK. Smartphone-Based Artificial Intelligence-Assisted Prediction for Eyelid Measurements: Algorithm Development and Observational Validation Study. JMIR Mhealth Uhealth. 2021;9(10). https://doi.org/10.2196/32444
51. Kuo MT, Hsu BWY, Lin YS, et al. Comparisons of deep learning algorithms for diagnosing bacterial keratitis via external eye photographs. Sci Rep. 2021;11(1). https://doi.org/10.1038/S41598-021-03572-6
52. Ka K, Lai H, Sze C, et al. Artificial intelligence and external photographs in ophthalmology: a systematic review. Hong Kong J Ophthalmol. 2024;28(1):19–32. https://doi.org/10.12809/HKJO-V28N1-384

Figures

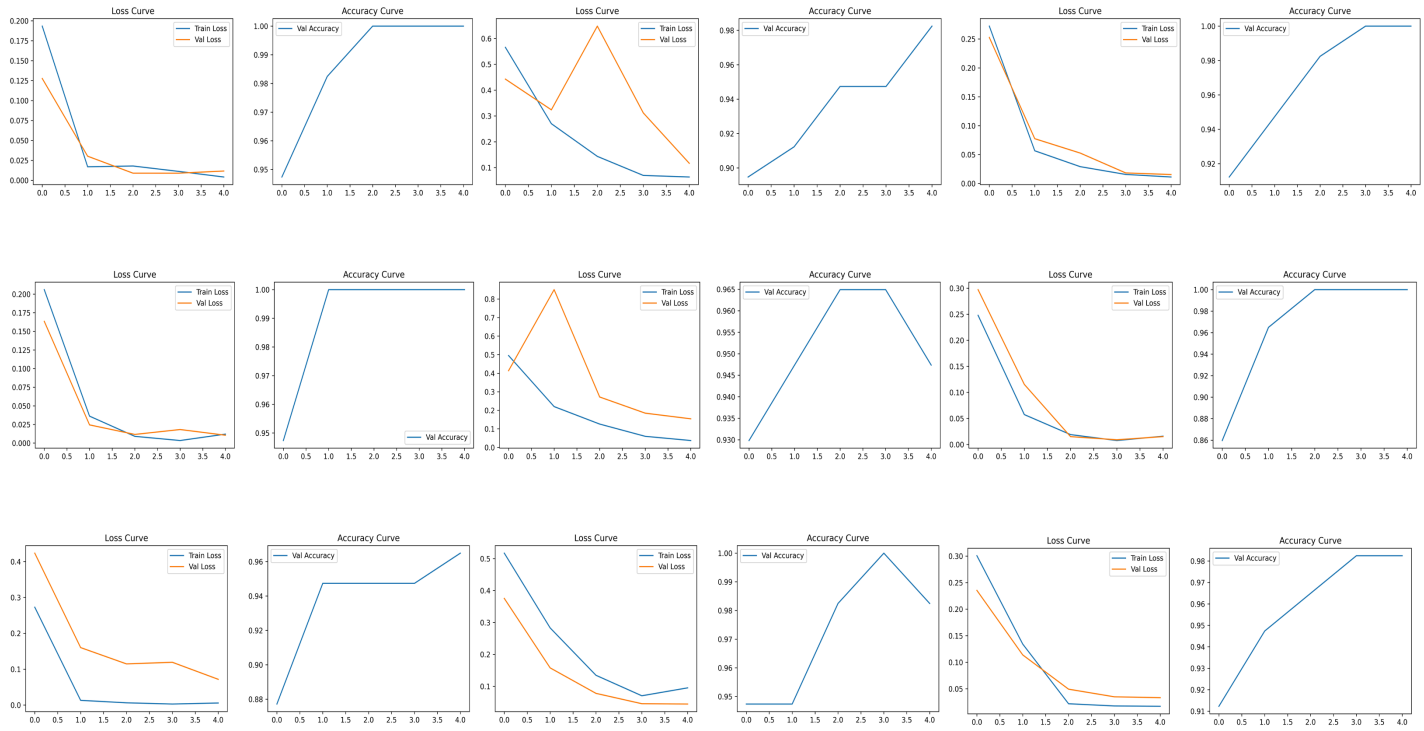


Figure 1

Training and validation curves for the three CNN architectures under all data-augmentation regimes. The first row shows the full_aug condition, the second row shows the no_blur condition, and the third row shows the no_noise condition.

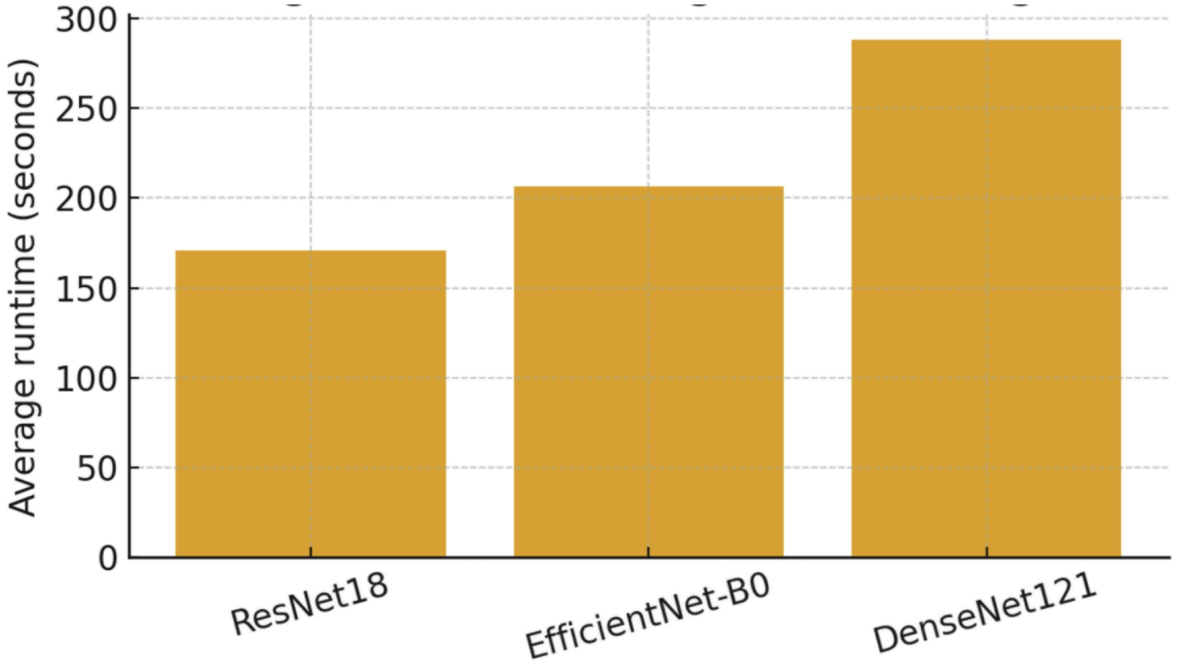


Figure 2

Comparison of the average combined training and evaluation runtime of ResNet18, EfficientNet-B0, and DenseNet121 across the three data-augmentation regimes.

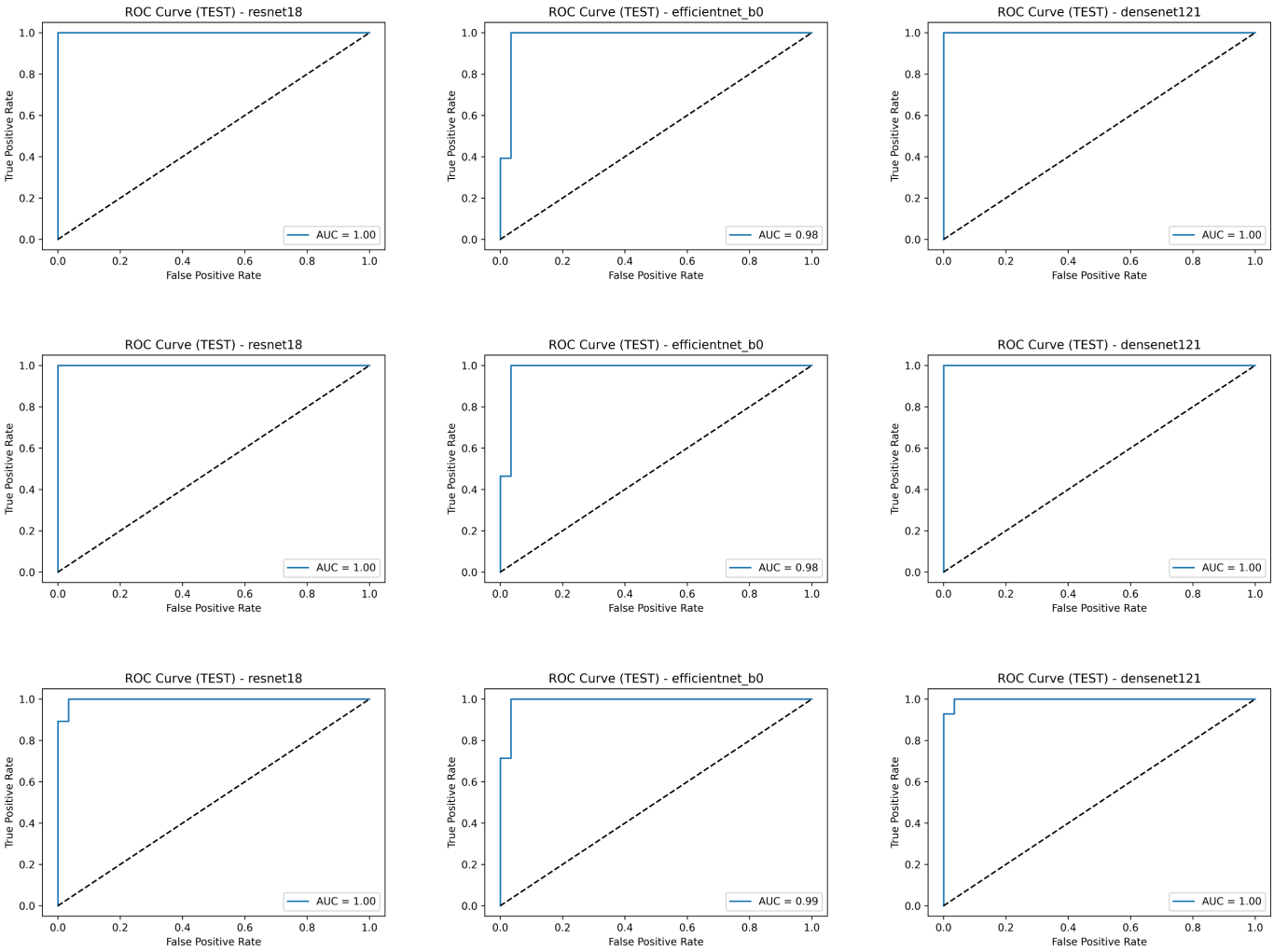


Figure 3

Receiver operating characteristic (ROC) curves for ResNet18, EfficientNet-B0, and DenseNet121 across the three data-augmentation regimes. Rows are arranged as follows: first row, full_aug; second row, no_blur; third row, no_noise.