

Supplementary Table 1. Summary of expert model parameters and training settings

Parameter	Value
Dataset information	
Number of training samples	960
Number of validation samples	240
Number of test samples	300
Class names	Dry AMD, normal, wet AMD
Mean (RGB channels)	0.2216, 0.2216, 0.2216
Standard deviation (RGB channels)	0.1876, 0.1876, 0.1876
Data augmentation techniques	
Image resize	224×224 pixels
Horizontal flip	Probability p=0.5
Random brightness contrast	Probability p=0.2
Normalization	Using computed mean and standard deviation
Model architecture	
Base model	EfficientNetV2 rw s (pretrained)
Classifier layers	- Linear layer: Input features → 512
	- Activation: LeakyReLU
	- Batch normalization: 512 units
	- Dropout layer: p=0.2821 (optimized)
	- Linear layer: 512 → 3 (number of classes)
Activation function	LeakyReLU
Weight initialization	He initialization (Kaiming normal)
Training settings	
Loss function	Cross-entropy loss with label smoothing ($\epsilon=0.1$)
Optimizer	AdamW
Learning rate	0.00138
Weight decay	0.00118
Number of epochs	25
Batch size	32
Learning rate	Scheduler StepLR (step size=7, $\gamma=0.1$)
Gradient clipping	Max norm=1.0
Mixed precision training	Enabled (using GradScaler)
L1 regularization	$\lambda L1 = 1 \times 10^{-5}$
AMD: Age-related macular degeneration	