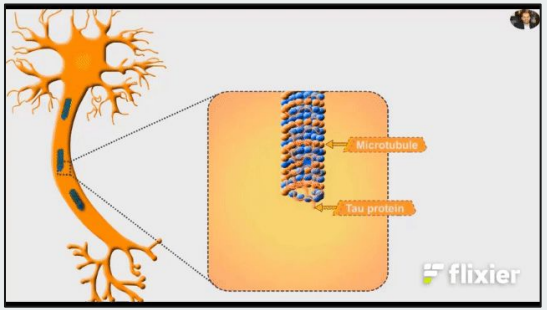
Tecnológico de Monterrey

CVPR Nashville JUNE 11-15, 2025

Beyond Neurofibrillary Tangles: Explainable AI for Microscopic Tauopathy Classification in Immunofluorescence Imaging

Jesus Dassaef López-Barrios, Miguel Angel Ontiveros-Torres, and Jose Antonio Cantoral-Ceballos

MSc CS & DS intern @ MSFT



June 12th, 2025

Why Tau Matters

📈 **Growing** Dementia Burden; Microscopy Remains **Manual** & **Late**

Why should we care about dementia?

World Health Organization

55

million people live with dementia

7th

leading cause of death

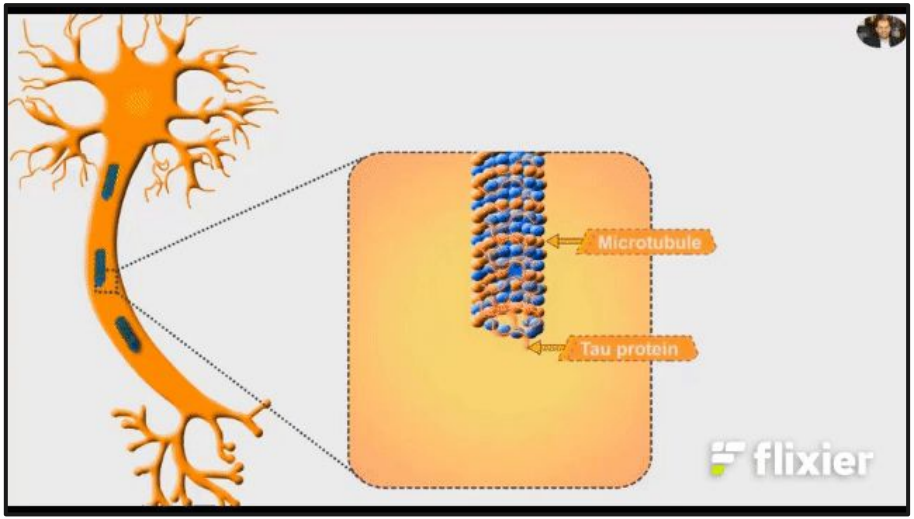
Caused by

diseases and injuries that affect the brain, such as Alzheimer's disease and stroke

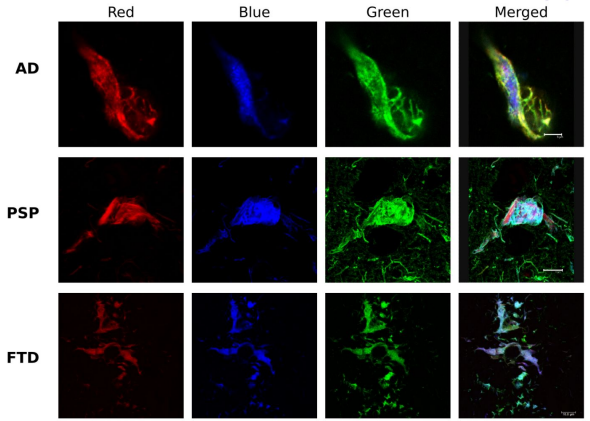
NFTs as a Diagnostic Filter

🕸 **Neurofibrillary Tangles** – **Powerful**, but **Not** the *Whole Picture?*

Tau misfolding → NFT formation



Actual dataset



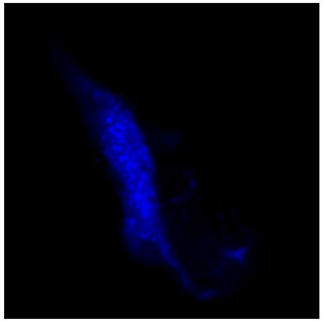
Mexican post-mortem hippocampal samples: triple-signal images (Alzheimer's Disease (AD), Progressive Supranuclear Palsy (PSP), and Frontotemporal Dementia (FTD)) with AT8-pTau (G), Thiazine-Red fibrils (R), pS396-Tau (B) + merged

Data: Rare & Carefully Controlled

2 150 Images • +10-Year Curation • Strict Bias Controls

- **Grayscale** conversion to suppress colour cues
- **512×512 px**, 100× oil objective → TIFF to NPY
- **Augmentation** confined to **train split**
- Dataset **National Dementia BioBank, Mexico;** shaped by **cultural donation barriers**

Blue

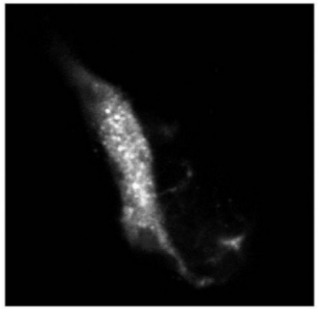


Dataset & Augmentation

1.1k images → 2.1k (grayscale, flips)

Split	AD	FTD	PSP	Total
Train	410	512	588	1510
Validation	86	108	126	320
Test	86	108	126	320
Total (Pre-Aug)	291	364	420	1075
Total (Aug)	582	728	840	2150

Grayscale



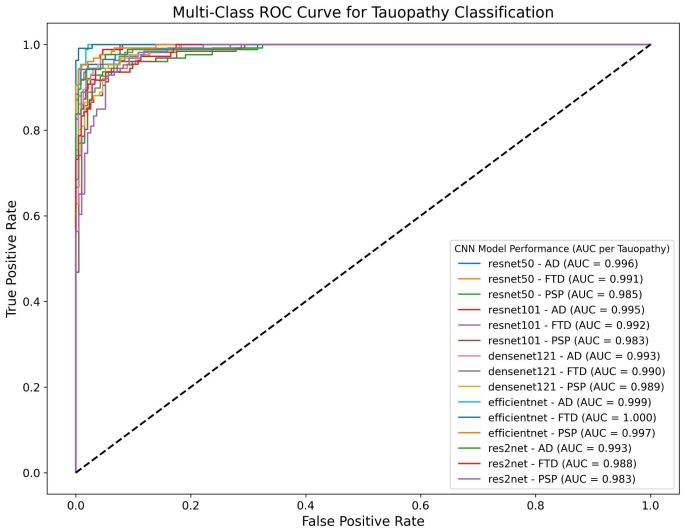
Model Setup & Baseline Metrics

Five CNNs – EfficientNet-B0 Leads, All >90 % F1

Model	Fine-Tuned Layers
ResNet50, ResNet101, Res2Net	Layers 1, 2, and 4
DenseNet121	Last 8 layers
EfficientNet-B0	Last 8 layers

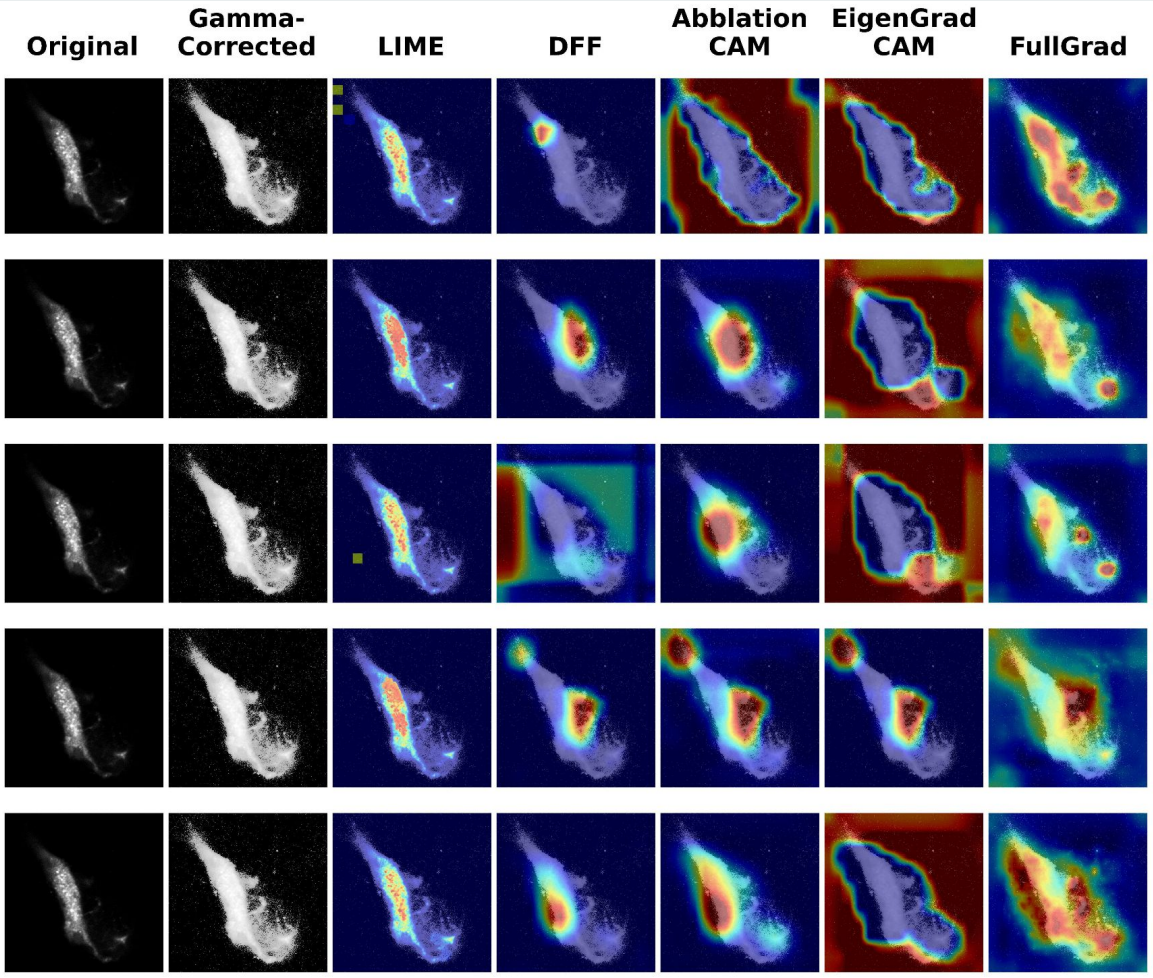
Hyperparameter	Value
Learning Rate	0.001
Batch Size	16
Optimizer	AdamW (weight decay $5e^{-4}$)
Loss Function	CrossEntropyLoss ($T_{max} = 10$, $\eta_{min} = 10^{-6}$)
Gradient Clipping	Max norm = 1.0
Scheduler	Cosine Annealing LR
Epochs	30

Model	Accuracy	Precision	Recall	F1-Score
EfficientNet-B0	**96.56**	**96.39**	**96.70**	**96.52**
ResNet101	92.50	92.51	92.56	92.53
DenseNet121	92.19	92.15	92.27	92.18
Res2Net	90.94	91.00	90.97	90.98
ResNet50	90.94	91.08	90.90	90.78

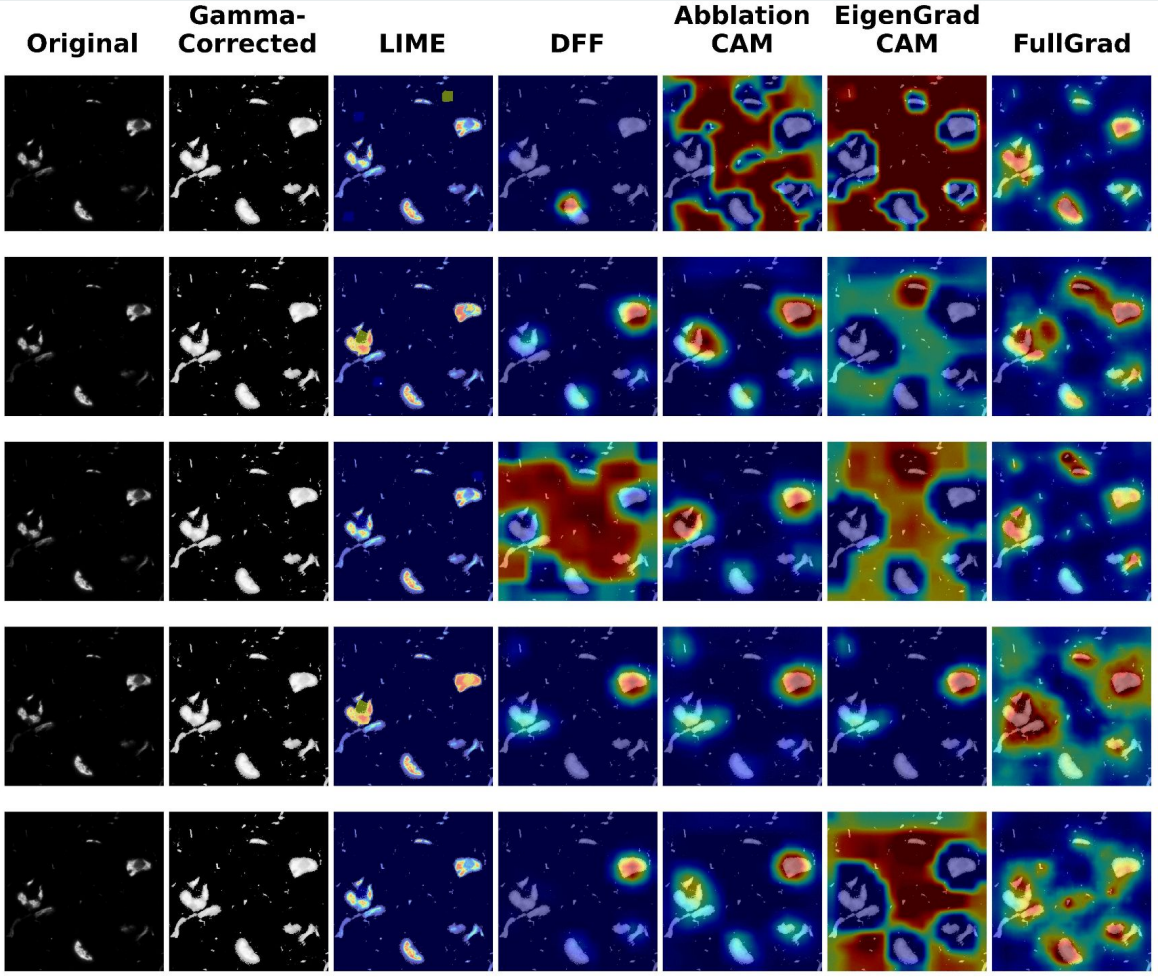


Explainable AI (XAI)

XAI grids


Original
Gamma-Corrected
LIME
DFF
Abblation CAM
EigenGrad CAM
FullGrad
EFFICIENTNET
RESNET101
DENSENET121
RES2NET
RESNET50

XAI grids


Original
Gamma-Corrected
LIME
DFF
Abblation CAM
EigenGrad CAM
FullGrad
EFFICIENTNET
RESNET101
DENSENET121
RES2NET
RESNET50

XAI Results

Model	EigenGradCAM	DFF	LIME	FullGrad	AblationCAM
ResNet50	Background	Tangle	Tangle	Tangle	Tangle
Res2Net	Tangle	Tangle	Tangle	Tangle	Tangle
DenseNet121	Background	**Background**	Tangle	Tangle	Tangle
ResNet101	Background	Tangle	Tangle	Tangle	Tangle
EfficientNet-B0	Background	Tangle	Tangle	Tangle	**Background**

Background (inverse effect) highlighted by **EigenGradCAM**, AblationCAM & DFF in high-F1 models; LIME/FullGrad stay **NFT-centric**

‘*Inverse Saliency*’: ***Background Highlights*** **Increase**  with **Performance**

Inverse-effect frequency bar

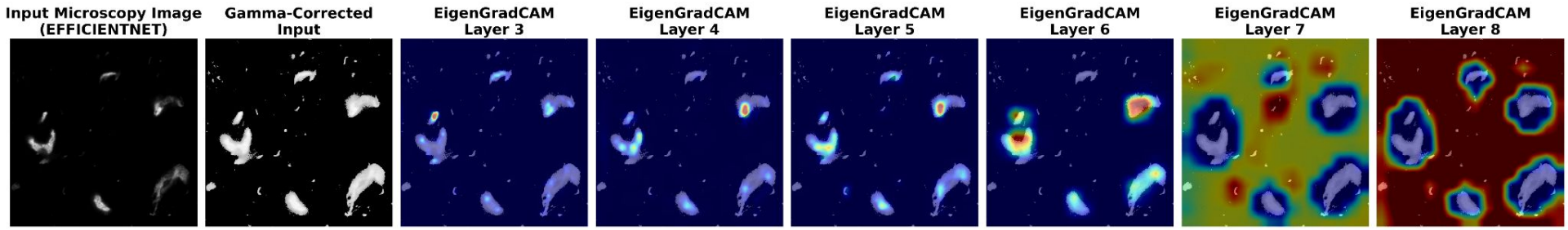
Model	Inverse Effect Frequency (%)
EfficientNet-B0	94.69
DenseNet121	74.69
ResNet101	72.81
ResNet50	68.75
Res2Net	38.75

Perturbation Test

👁️🧱 *Occluding* **Background** *Alters* Predictions – But *Interpretation* ≠ ***Causation***

Layer-wise saliency

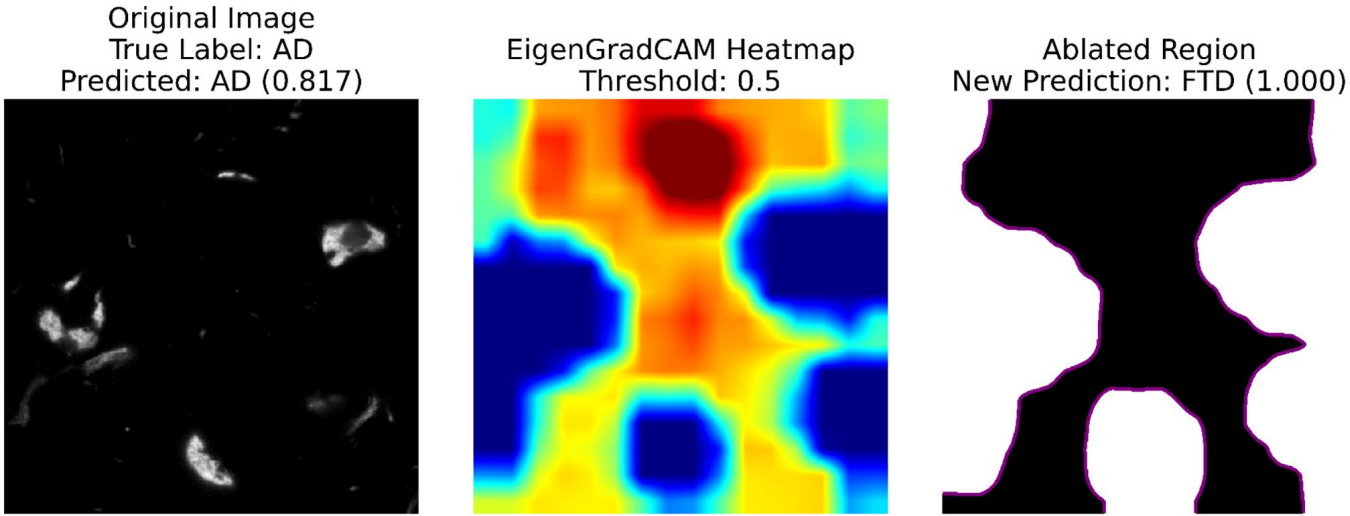
- EigenGradCAM across EfficientNet layers
- Granular to background
- Learnt not spontaneous



Ablation metrics

Model	% Inverse Effect (EigenGradCAM)	% No Inverse Effect (EigenGradCAM)	Important Feature (if removed)	Inverse	avg_flip_rate (% class change)	avg_delta_p (decrease in true class score)	More Important
DenseNet121	**75%**	25%	**Background**	**No**	**81%**	**0.72**	**Background**
	25%	75%	Tangle	Yes	68%	0.58	
EfficientNet	**95%**	5%	**Background**	**No**	**65%**	**0.61**	**Both**
	5%	95%	Tangle	**Yes**	**65%**	**0.63**	
ResNet101	**73%**	27%	**Background**	**No**	**61%**	**0.53**	**Both**
	27%	73%	Tangle	**Yes**	**61%**	**0.53**	
Res2Net	39%	61%	**Tangle**	**No**	**66%**	**0.57**	**Both**
	61%	39%	Background	**Yes**	**66%**	**0.57**	
ResNet50	**69%**	31%	**Background**	No	53%	0.47	**Tangle**
	31%	69%	Tangle	**Yes**	**70%**	**0.57**	

Example triple-panel



Conclusions & Next Steps

Beyond NFTs → Non-Canonical Tauopathy Signals

- **Solid but limited accuracy –** EfficientNet-B0 tops; all 5 CNNs > 90 % on every metric.
- **Saliency –** High-F1 models spotlight *non-canonical* background rather than tangles.
- **Ablation –** Removing those zones flips labels (e.g., AD → FTD) / drops confidence ≈ NFT removal.
- ***Interpretation** – May hint at neuropil threads or glial inclusions or spurious bias; expert confirmation essential.*
- **Next steps –** Multi-site cohorts + GFAP / IBA1 stains, multimodal PET/MRI, stronger bias tests.
- **Clinical role –** AI proposes hypotheses; neuropathologists decide.

Background features: **potential biomarkers or artifacts**? **Validation** will tell.

Future work: Initial base *encoder* for full *captioning* report generation explainable pipeline.

Thanks!

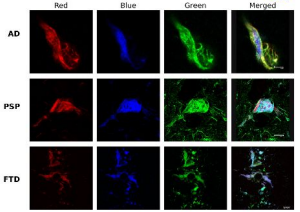
Tecnológico de Monterrey

Beyond Neurofibrillary Tangles: Explainable AI for Microscopic Tauopathy Classification in Immunofluorescence Imaging

Jesus Dassaef López-Barrios, Miguel Angel Ontiveros-Torres, and Jose Antonio Cantoral-Ceballos

CVPR Nashville

Related Work & Methodology



Mexican post-mortem hippocampal samples: triple-signal images (Alzheimer's Disease (AD), Progressive Supranuclear Palsy (PSP), and Frontotemporal Dementia (FTD)) with AT8-pTau (G), Thiazine-Red fibrils (R), pS396-Tau (B) + merged

Dataset & Augmentation

1.1k images → 2.1k (grayscale, flips)

Split	AD	FTD	PSP	Total
Train	410	512	588	1510
Validation	86	108	126	320
Test	86	108	126	320
Total (Pre-Aug)	291	364	420	1075
Total (Aug)	582	728	840	2150

Architectures

Model	Fine-Tuned Layers
ResNet50, ResNet101, Res2Net	Layers 1, 2, and 4
DenseNet121	Last 8 layers
EfficientNet-B0	Last 8 layers

Head

Layer	Neurons	Regularization
FC 1	1024	BatchNorm + Dropout (0.6)
FC 2	512	BatchNorm + Dropout (0.5)
Output	3	Softmax

Results

Model Performance

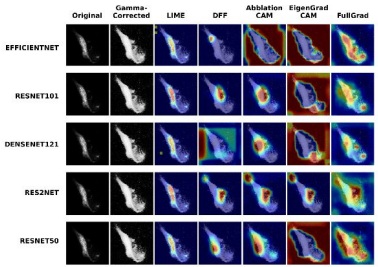
Top model = EfficientNet-B0; all 5 CNNs **exceed 90%** on every metric

Model	Accuracy	Precision	Recall	F1-Score
EfficientNet-B0	**96.56**	**96.39**	**96.70**	**96.52**
ResNet101	92.50	92.51	92.56	92.53
DenseNet121	92.19	92.15	92.27	92.18
Res2Net	90.94	91.00	90.97	90.98
ResNet50	90.94	91.08	90.90	90.78

XAI Results

Model	EigenGradCAM	DFF	LIME	FullGrad	AblationCAM
ResNet50	Background	Tangle	Tangle	Tangle	Tangle
Res2Net	Tangle	Tangle	Tangle	Tangle	Tangle
DenseNet121	Background	**Background**	Tangle	Tangle	Tangle
ResNet101	Background	Tangle	Tangle	Tangle	Tangle
EfficientNet-B0	Background	Tangle	Tangle	Tangle	**Background**

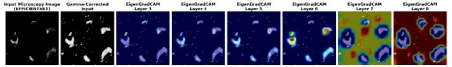
Background (inverse effect) highlighted by **EigenGradCAM**, AblationCAM & DFF in high-F1 models; LIME/FullGrad stay **NFT-centric**



EigenGradCAM Effects Across Models and Layers

Model	Inverse Effect Frequency (%)
EfficientNet-B0	94.69
DenseNet121	74.69
ResNet101	72.81
ResNet50	68.75
Res2Net	38.75

Inverse saliency frequency rises with performance

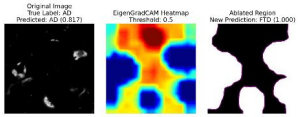


Layer progression. EfficientNet example: early layers focus on tangles, deep layers shift to diffuse background

Ablation Study

Model	% Inverse Effect (EigenGradCAM)	% No Inverse Effect (EigenGradCAM)	Important Feature (if removed)	Inverse	avg_flip_rate (% class change)	avg_delta_p (decrease in true class score)	More Important
DenseNet121	75%	25%	**Background**	No	81%	0.72	**Background**
	25%	75%	Tangle	Yes	68%	0.58	
EfficientNet	95%	5%	**Background**	No	65%	0.61	**Both**
	5%	95%	Tangle	Yes	65%	0.63	
ResNet101	73%	27%	**Background**	No	61%	0.53	**Both**
	27%	73%	Tangle	Yes	61%	0.53	
Res2Net	39%	61%	**Tangle**	No	66%	0.57	**Both**
	61%	39%	Background	Yes	66%	0.57	
ResNet50	69%	31%	**Background**	No	53%	0.47	**Tangle**
	31%	69%	Tangle	Yes	70%	0.57	

Ablation takeaway: DenseNet121 is background-driven; EfficientNet-B0, ResNet101 & Res2Net rely on both regions; ResNet50 remains tangle-centric



Ablation exemplar. Removing EigenGradCAM-derived background flips prediction AD → FTD, underscoring non-tangle cues

Biological Insights & Hypotheses

- Saliency suggests CNNs may detect **pathology beyond NFTs**, such as neuropil threads or glial inclusions.
- Background attribution may reflect **relevant features or artifacts**; further validation needed.
- Supports using **additional markers and datasets** to verify non-canonical Tau patterns.