Early Diagnosis of Lyme Disease by Recognizing Erythema Migrans Skin Lesion from Images Utilizing Deep Learning Techniques

Sk Imran Hossain, Engelbert Mephu Nguifo, Jocelyn de Goër de Herve

Early diagnosis of Lyme disease by recognizing Erythema Migrans skin lesion from images utilizing deep learning techniques

International Conference on Data Science and Advanced Analytics (DSAA) 2021
Doctoral Consortium

PREPARED BY

Sk Imran HOSSAIN, PhD Student, LIMOS, UCA

sk_imran.hossain@uca.fr

SUPERVISED BY

Engelbert MEPHU NGUIFO

Engelbert.MEPHU_NGUIFO@uca.fr

CO-SUPERVISED BY

Jocelyn DE GOËR

jocelyn.degoer@inrae.fr

FUNDING : Région : projet DAPPEM–AV0021029 + FEDER

"The DAPPEM project was carried out under the Call for Proposal "Pack Ambition Research" from the Auvergne-Rhône-Alpes region, France"

Presentation Outline

- Research Problem
- Research Progress
- Research Plan

Research Problem
Lyme disease

Tick

Source

Erythema Migrans (EM)

Bull's-eye pattern

Source

Atypical pattern

Source

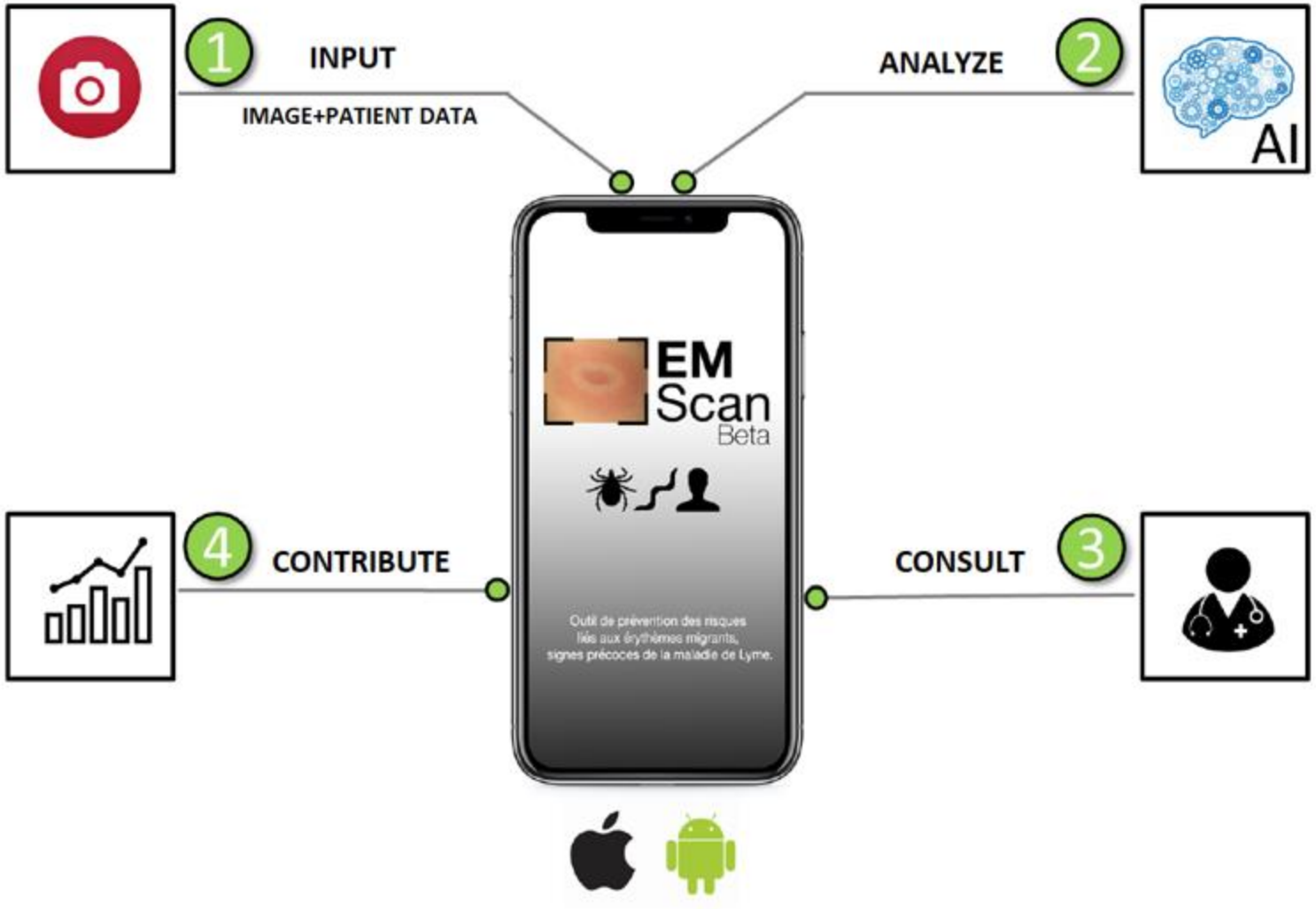
Complications

- Facial nerve paralysis
- Heart Palpitations
- Arthritis
- Meningitis

Frequency

- 50000 per year in France
- 300000 per year in USA

Research Problem
DAPPEM project goal



Research Problem
Related works

1. Čuk, E., et al., 2014. Supervised visual system for recognition of erythema migrans, an early skin manifestation of lyme borreliosis. Stroj. Vestnik/Journal Mech. Eng. 60, 115–123. https://doi.org/10.5545/sv-jme.2013.1046

Classical machine learning techniques
(Private Dataset)

2. Burlina, P., et al., 2018. Skin image analysis for erythema migrans detection and automated lyme disease referral, in: LNCS (Including Subseries LNBI). pp. 244–251. https://doi.org/10.1007/978-3-030-01201-4_26

ResNet50 [1] Deep Convolutional Neural Network (DCNN)
(Private Dataset)

Limitations: considers image data only

Research Problem
Problem statement

Input Image

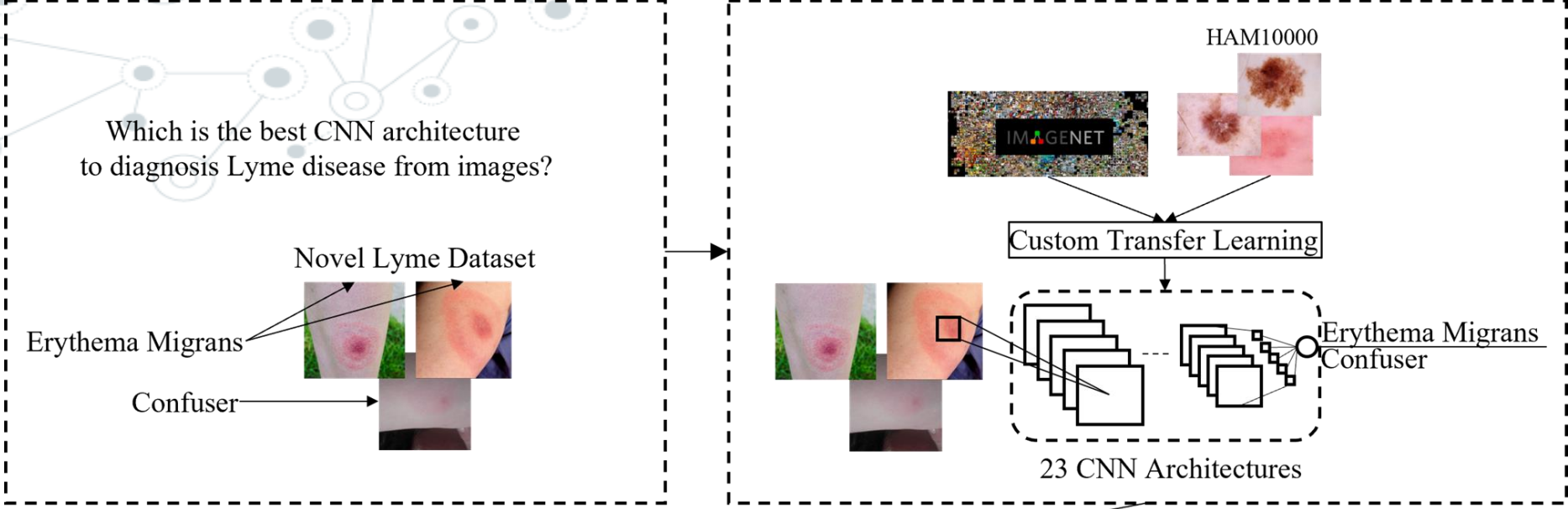
Patient Profile

Age
Recent Forest Visit?
.....
.....

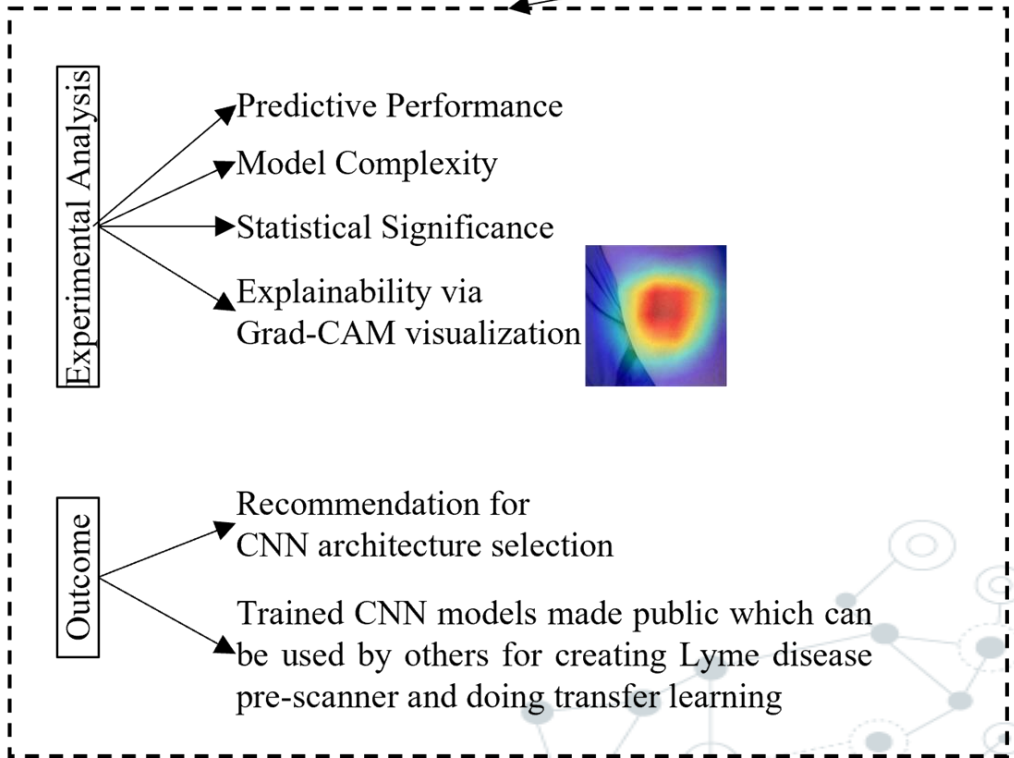
Prediction Model

Prediction

Research Progress
Benchmarking DCNN architectures



Recent studies show that deep learning methods compete on par with expert dermatologists to identify skin lesions from the image [2]



Research Progress
Benchmarking DCNN architectures (Dataset preparation)

CLERMONT FERRAND

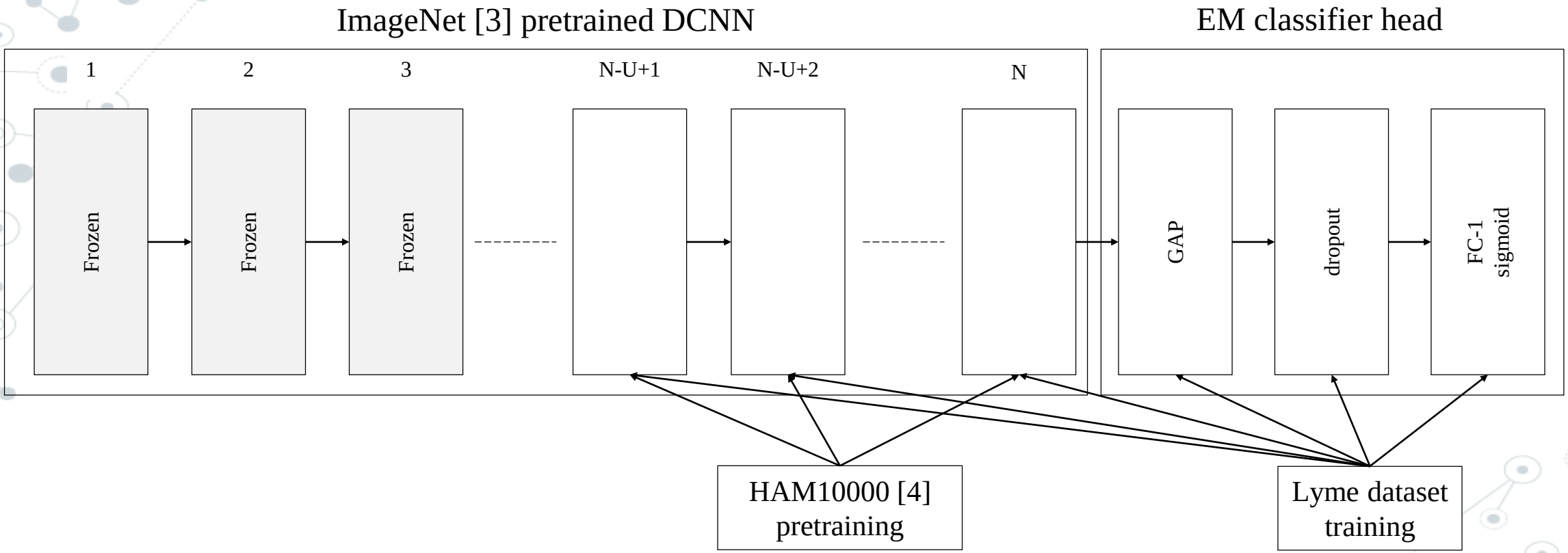
Total Images : 1672

Binary Classification

Erythema Migrans (EM)
866/1672 (51.8%)

Confuser
806/1672 (48.2%)

Research Progress
Benchmarking DCNN architectures (Transfer learning)



Research Progress
Experimental results: Burlina model

Table 1: Five-fold cross-validation performance metrics of Burlina ResNet50 based model. Within each cell, the value after (±) symbol represents the standard deviation across five folds.

Metric / Model	*Accuracy*	*Sensitivity*	*Specificity*	*Precision*	*NPV*	*MCC*	*Kappa*	*LR+*	*LR-*	*F1-Score*	*AUC*
ResNet50-Burlina	76.05 ±0.74	70.05 ±3.6	82.51 ±3.31	81.29 ±2.1	72.04 ±1.71	0.5294 ±0.0132	0.5229 ±0.0145	4.1017 ±0.5172	0.362 ±0.0309	0.7515 ±0.0137	0.481 ±0.0509

Research Progress

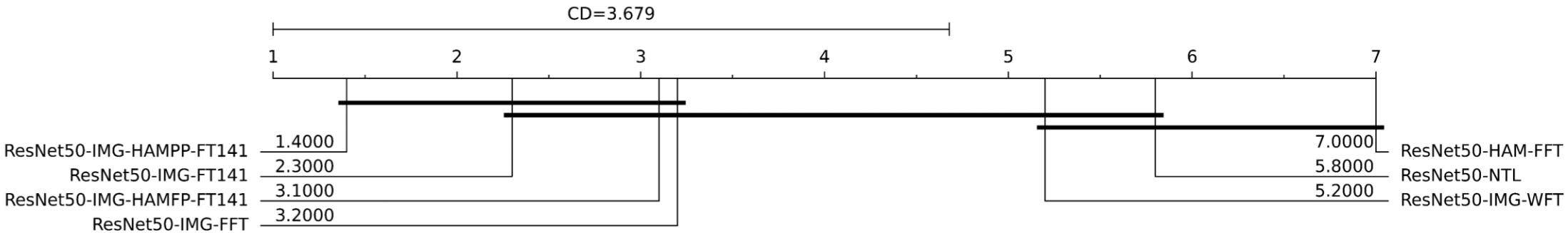
Experimental results: experiments with ResNet50 architecture

Metric / Model	Accuracy	Sensitivity	Specificity	Precision	NPV	MCC	Kappa	LR+	LR-	F1-Score	AUC
ResNet50-NTL	76.35 ±2.43	78.49 ±8.47	74.04 ±4.6	76.64 ±1.64	76.92 ±5.22	0.5305 ±0.0431	0.5261 ±0.0464	3.0735 ±0.2867	0.2853 ±0.0906	0.7723 ±0.0398	0.8471 ±0.0185
ResNet50-HAM-FFT	72.27 ±1.69	75.85 ±1.27	68.42 ±4.05	72.18 ±2.55	72.48 ±1.08	0.4447 ±0.0341	0.4435 ±0.0347	2.4434 ±0.3248	0.3536 ±0.0193	0.7393 ±0.0116	0.7979 ±0.0251
ResNet50-IMG-WFT	78.94 ±1.48	82.55 ±2.77	75.06 ±5.11	78.27 ±3.2	80.11 ±1.77	0.5799 ±0.03	0.5772 ±0.0305	3.4636 ±0.7671	0.2316 ±0.0255	0.8025 ±0.0101	0.8666 ±0.0163
ResNet50-IMG-FFT	82.22 ±1.36	85.27 ±2.67	78.93 ±5.26	81.55 ±3.42	83.42 ±1.63	0.6458 ±0.0262	0.6431 ±0.028	4.3127 ±1.0994	0.1854 ±0.0226	0.8326 ±0.0083	0.909 ±0.0092
ResNet50-IMG-FT141	83.24 ±1.04	85.29 ±2.27	**81.04** ±2.28	82.91 ±1.49	83.74 ±1.96	0.6649 ±0.0212	0.6641 ±0.021	4.5575 ±0.493	0.1812 ±0.0255	0.8405 ±0.0104	0.9134 ±0.0091
ResNet50-IMG-HAMFP-FT141	82.35 ±1.62	**89.28** ±2.42	74.91 ±5.11	79.45 ±3.05	**86.81** ±2.03	0.6521 ±0.0295	0.6448 ±0.0333	3.7072 ±0.7368	**0.1421** ±0.0251	0.84 ±0.0111	0.9113 ±0.0091
ResNet50-IMG-HAMPP-FT141	**84.42** ±1.36	87.93 ±1.47	80.65 ±3.59	**83.1** ±2.49	86.19 ±1.27	**0.6893** ±0.0263	**0.6874** ±0.0277	**4.703** ±0.8624	0.1493 ±0.0155	**0.8541** ±0.0106	**0.9189** ±0.0115

Research Progress

Experimental results: experiments with ResNet50 architecture

CD Diagram on Accuracy



Friedman test p value of 0.0002

Research Progress

Experimental results: performance metrics of trained DCNN architectures

Metric / Model	Accuracy	Sensitivity	Specificity	Precision	NPV	MCC	Kappa	LR+	LR-	F1-Score	AUC
VGG16-8	82.17 ±1.23	85.77 ±3.58	78.31 ±4.36	81.12 ±2.62	83.88 ±3.02	0.6453 ±0.0253	0.6422 ±0.0249	4.0983 ±0.7329	0.1802 ±0.0388	0.8328 ±0.0116	0.9011 ±0.0079
VGG19-13	84.14 ±1.62	85.29 ±1.69	82.9 ±2.63	**84.32** ±1.97	84.0 ±1.67	0.6826 ±0.0323	0.6823 ±0.0326	5.0924 ±0.6884	0.1777 ±0.0214	0.8479 ±0.0146	0.913 ±0.0074
ResNet50-141	**84.42** ±1.36	87.93 ±1.47	80.65 ±3.59	83.1 ±2.49	86.19 ±1.27	**0.6893** ±0.0263	**0.6874** ±0.0277	4.703 ±0.8624	0.1493 ±0.0155	**0.8541** ±0.0106	**0.9189** ±0.0115
ResNet101-150	82.64 ±2.1	83.68 ±3.49	81.52 ±2.29	82.97 ±1.8	82.4 ±3.09	0.6528 ±0.0419	0.6522 ±0.0418	4.6001 ±0.6257	0.2004 ±0.0427	0.8329 ±0.022	0.9044 ±0.0094
ResNet50V2-105	82.37 ±2.15	85.53 ±3.35	78.96 ±6.13	81.66 ±3.83	83.72 ±2.63	0.6493 ±0.0411	0.6461 ±0.0439	4.3618 ±1.0495	0.1819 ±0.0349	0.8343 ±0.017	0.9013 ±0.0133
ResNet101V2-233	82.58 ±2.21	81.9 ±4.78	**83.32** ±3.71	84.17 ±2.55	81.31 ±3.7	0.6535 ±0.0429	0.6515 ±0.0439	**5.104** ±0.9811	0.2163 ±0.0541	0.8292 ±0.0254	0.9118 ±0.0149
InceptionV3-274	82.73 ±2.08	86.57 ±2.42	78.6 ±2.8	81.33 ±2.12	84.52 ±2.52	0.6551 ±0.0419	0.6533 ±0.0419	4.1259 ±0.639	0.1714 ±0.0328	0.8385 ±0.0195	0.9052 ±0.0185
InceptionV4-327	82.76 ±1.78	85.7 ±3.96	79.58 ±2.87	81.92 ±1.8	84.02 ±3.41	0.6561 ±0.0358	0.6541 ±0.0353	4.2716 ±0.5734	0.179 ±0.0465	0.837 ±0.0197	0.9092 ±0.019
InceptionResNetV2-500	82.67 ±2.06	83.54 ±3.88	81.74 ±3.16	83.17 ±2.16	82.37 ±3.32	0.6541 ±0.0406	0.653 ±0.041	4.6886 ±0.7264	0.2009 ±0.0456	0.8329 ±0.0218	0.9011 ±0.0133
Xception-118	82.48 ±2.45	83.16 ±5.1	81.75 ±2.76	83.08 ±1.94	82.16 ±4.4	0.6507 ±0.0487	0.6492 ±0.0484	4.6434 ±0.6571	0.2054 ±0.0609	0.8304 ±0.0276	0.9081 ±0.0148

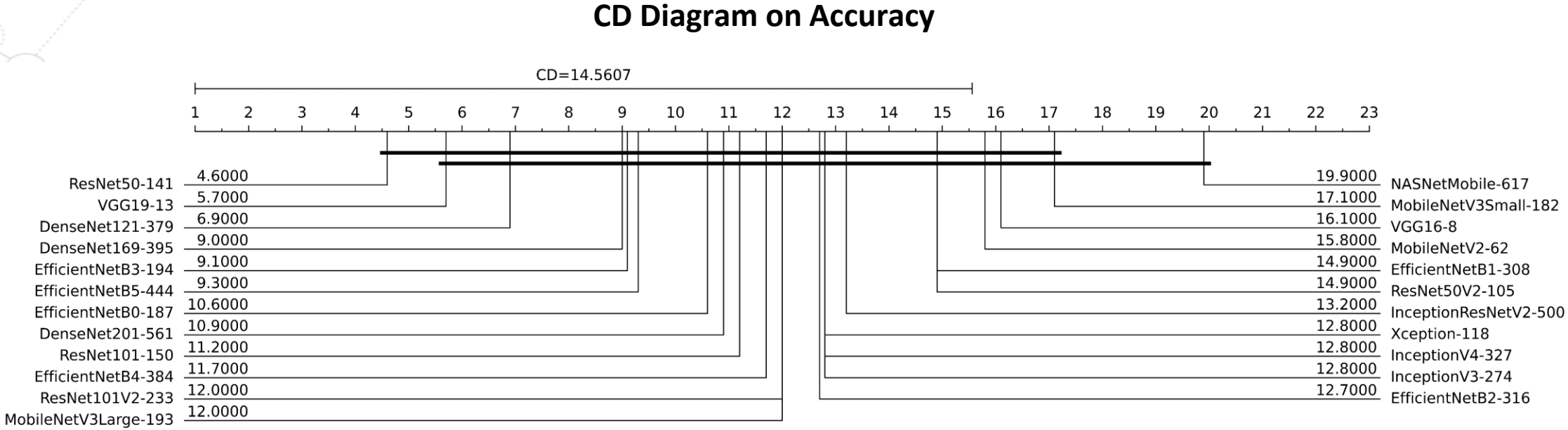
Benchmarking

Experimental results: performance metrics of trained DCNN architectures

Metric / Model	*Accuracy*	*Sensitivity*	*Specificity*	*Precision*	*NPV*	*MCC*	*Kappa*	*LR+*	*LR-*	*F1-Score*	*AUC*
DenseNet121-379	83.88 ±0.92	85.85 ±1.76	81.75 ±0.95	83.49 ±0.69	84.35 ±1.57	0.6773 ±0.0186	0.6768 ±0.0184	4.7169 ±0.254	0.173 ±0.0211	0.8465 ±0.01	0.9158 ±0.0097
DenseNet169-395	83.66 ±1.25	**88.6** ±3.59	78.35 ±2.75	81.54 ±1.53	**86.68** ±3.33	0.6758 ±0.0265	0.6717 ±0.0249	4.1454 ±0.4187	**0.1446** ±0.0414	0.8486 ±0.0138	0.9123 ±0.0129
DenseNet201-561	83.12 ±1.11	85.61 ±1.81	80.45 ±3.92	82.61 ±2.7	83.93 ±1.13	0.663 ±0.0221	0.6615 ±0.0228	4.5729 ±0.9885	0.1783 ±0.0153	0.8403 ±0.0073	0.9125 ±0.0083
MobileNetV2-62	81.68 ±1.99	81.94 ±3.49	81.39 ±1.26	82.55 ±1.21	80.85 ±3.09	0.6337 ±0.0394	0.6332 ±0.0395	4.4256 ±0.3596	0.222 ±0.0441	0.8222 ±0.0218	0.8933 ±0.0135
MobileNetV3Small -182	81.53 ±1.98	84.93 ±3.29	77.87 ±3.89	80.6 ±2.55	82.91 ±2.85	0.6315 ±0.0398	0.6294 ±0.04	3.9496 ±0.6356	0.1933 ±0.0386	0.8265 ±0.0186	0.896 ±0.013
MobileNetV3Large -193	82.74 ±2.17	83.69 ±0.43	81.71 ±4.6	83.26 ±3.39	82.3 ±0.89	0.6548 ±0.0437	0.6542 ±0.0442	4.8573 ±1.1585	0.2002 ±0.0117	0.8344 ±0.017	0.9034 ±0.0094
NASNetMobile - 617	81.3 ±1.45	83.2 ±1.66	79.25 ±3.98	81.29 ±2.65	81.48 ±1.07	0.6261 ±0.0287	0.6251 ±0.0297	4.1452 ±0.7283	0.2117 ±0.0156	0.8219 ±0.0108	0.8897 ±0.0152
EfficientNetB0- 187	83.13 ±1.2	85.21 ±3.91	80.89 ±2.95	82.83 ±1.75	83.79 ±3.19	0.6636 ±0.0244	0.6618 ±0.0237	4.5522 ±0.6116	0.1817 ±0.0427	0.8392 ±0.0147	0.9094 ±0.0129
EfficientNetB1- 308	82.42 ±1.04	85.85 ±2.14	78.71 ±3.75	81.37 ±2.34	83.9 ±1.59	0.6492 ±0.0202	0.647 ±0.0214	4.1494 ±0.6707	0.179 ±0.0209	0.835 ±0.0074	0.9088 ±0.0134
EfficientNetB2- 316	82.75 ±1.4	84.95 ±3.41	80.39 ±3.02	82.39 ±1.91	83.4 ±2.69	0.6556 ±0.0276	0.6542 ±0.0279	4.4211 ±0.6202	0.1865 ±0.0379	0.8359 ±0.0158	0.9075 ±0.0082
EfficientNetB3- 194	83.46 ±0.87	85.15 ±4.28	81.64 ±2.9	83.4 ±1.6	83.9 ±3.14	0.6704 ±0.0157	0.6685 ±0.0167	4.7361 ±0.6283	0.1803 ±0.0443	0.8416 ±0.0144	0.9163 ±0.0074
EfficientNetB4- 384	82.98 ±1.31	87.55 ±2.2	78.06 ±3.76	81.2 ±2.33	85.46 ±1.76	0.6613 ±0.0249	0.6581 ±0.0268	4.0946 ±0.6159	0.1589 ±0.0233	0.842 ±0.0107	0.9138 ±0.0074
EfficientNetB5- 444	83.7 ±1.21	86.85 ±2.89	80.32 ±3.73	82.71 ±2.39	85.17 ±2.38	0.6752 ±0.024	0.6729 ±0.0245	4.5562 ±0.7645	0.1629 ±0.0303	0.8466 ±0.0108	0.9138 ±0.0161

Research Progress

Experimental results: performance metrics of trained DCNN architectures



Friedman test p value of 0.0822

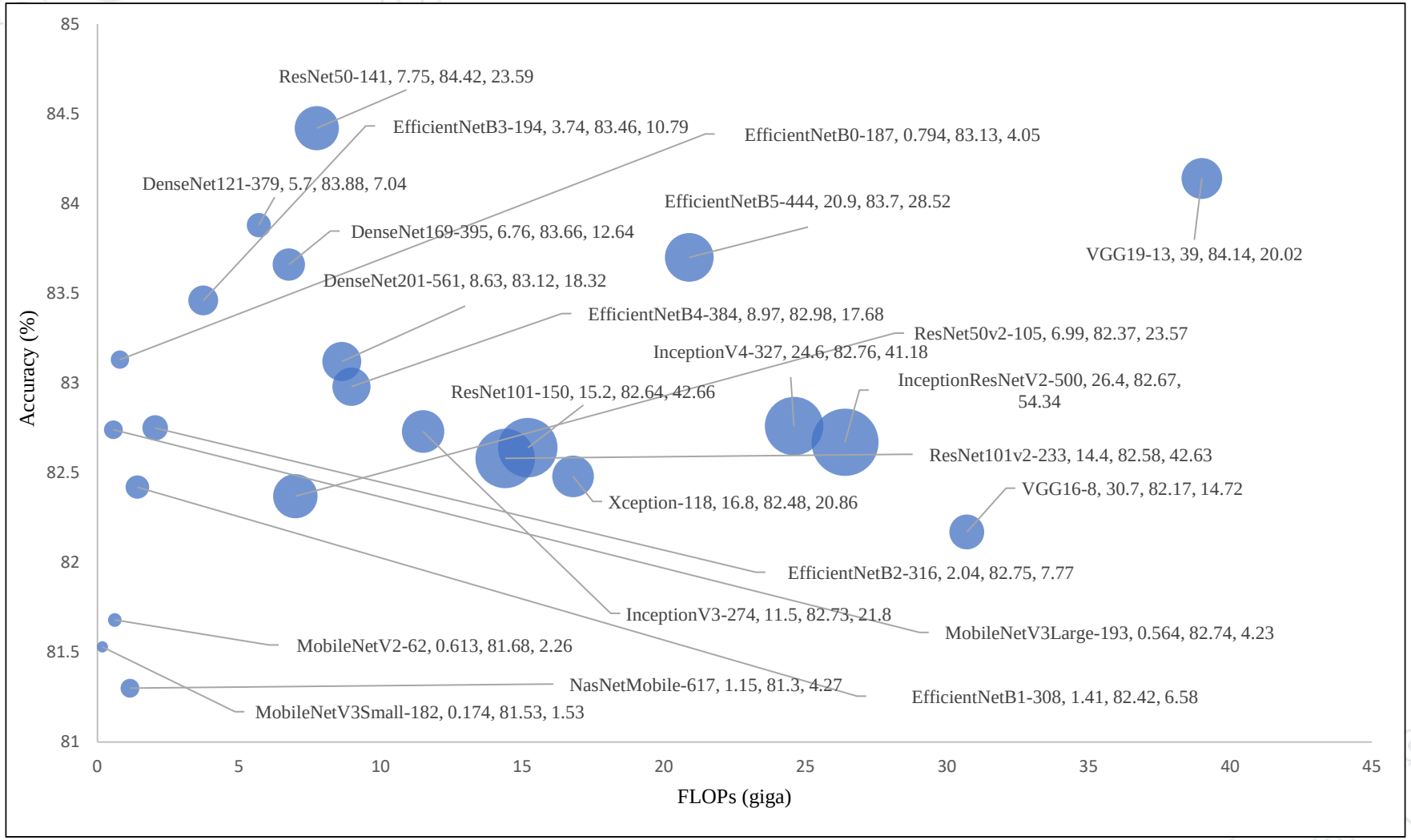
Research Progress

Experimental results: complexity metrics of trained DCNN architectures

Model	Parameters (million)	FLOPs (giga)	Average training time (sec per epoch)	Disk usage (megabyte)	GPU usage (megabyte)	Average inference time (sec per image)	Input shape
VGG16-8	14.72	30.7	111	146.24	565	0.0426	224x224x3
VGG19-13	20.02	39	164	216.03	565	0.0431	224x224x3
ResNet50-141	23.59	7.75	113	268.64	821	0.0484	224x224x3
ResNet101-150	42.66	15.2	123.33	378.41	821	0.0539	224x224x3
ResNet50V2-105	23.57	6.99	**76**	258.84	821	0.0464	224x224x3
ResNet101V2-233	42.63	14.4	152	429.45	821	0.0599	224x224x3
InceptionV3-274	21.8	11.5	133	246.94	821	0.0540	224x224x3
InceptionV4-327	41.18	24.6	223.33	424.93	1333	0.0735	299x299x3
InceptionResNetV2-500	54.34	26.4	281.33	588.12	1333	0.0958	299x299x3
Xception-118	20.86	16.8	243.33	238.48	821	**0.0392**	299x299x3
DenseNet121-379	7.04	5.7	140.67	78.7	437	0.0673	224x224x3
DenseNet169-395	12.64	6.76	130	128.59	565	0.0686	224x224x3
DenseNet201-561	18.32	8.63	182.67	198.7	565	0.0840	224x224x3
MobileNetV2-62	2.26	0.613	78	24.02	**341**	0.0429	224x224x3
MobileNetV3Small-182	**1.53**	**0.174**	81	**17.8**	**341**	0.0444	224x224x3
MobileNetV3Large-193	4.23	0.564	86.33	48.34	373	0.0444	224x224x3
NASNetMobile -617	4.27	1.15	152	50.74	373	0.0741	224x224x3
EfficientNetB0-187	4.05	0.794	87	46.59	373	0.0523	224x224x3
EfficientNetB1-308	6.58	1.41	158.33	75.89	437	0.0546	240x240x3
EfficientNetB2-316	7.77	2.04	210	89.53	437	0.0565	260x260x3
EfficientNetB3-194	10.79	3.74	143	117.6	565	0.0648	300x300x3
EfficientNetB4-384	17.68	8.97	431	202.2	565	0.0614	380x380x3
EfficientNetB5-444	28.52	20.9	771	325.21	821	0.0659	456x456x3

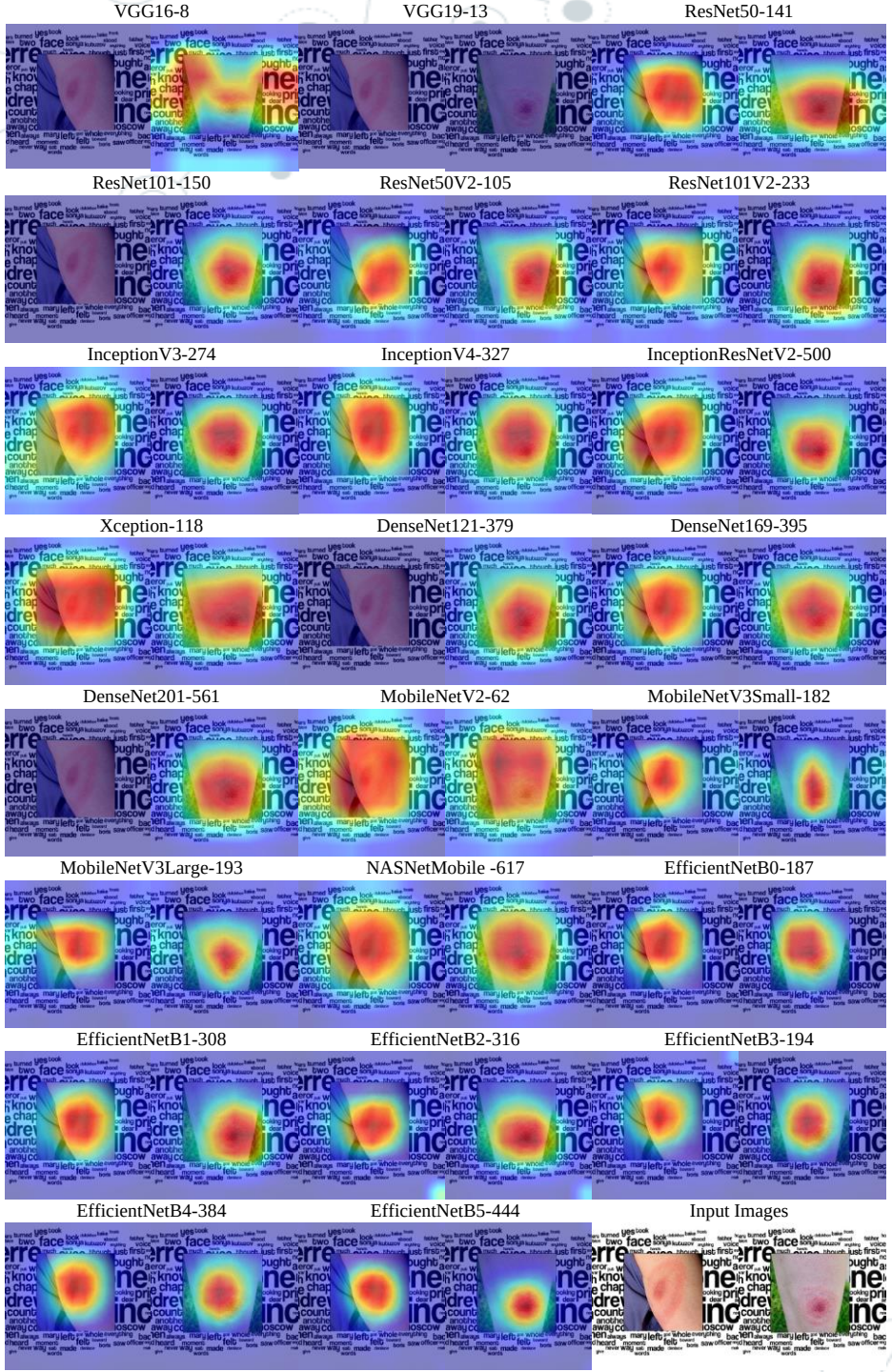
Research Progress

Experimental results: accuracy vs complexity bubble chart



Bubble chart reporting model accuracy vs floating-point operations (FLOPs). The size of each bubble represents number of model parameters measured in millions unit. Beside each model name the three values represent FLOPs, accuracy, and model parameters, respectively.

Research Progress
Experimental results: Grad-CAM Visualization



Research Progress

Experimental results: detailed results and trained models

- Publicly available at https://drive.uca.fr/d/f360e65bfcd84b59b7c0/

- Can be used by others for transfer learning, and building pre-scanners for Lyme diesease

Research Progress
Publication

- *"Exploring convolutional neural networks with transfer learning for diagnosing Lyme disease from skin lesion images"*
  - In revision - *Computer Methods and Programs in Biomedicine* journal
  - S. I. Hossain *et al.*, "Benchmarking convolutional neural networks for diagnosing Lyme disease from images," Jun. 2021. ⟨hal-03271706⟩

Research Plan
Skin hair augmentation








Research Plan
Skin hair augmentation

Skin Hair: Image processing based removal

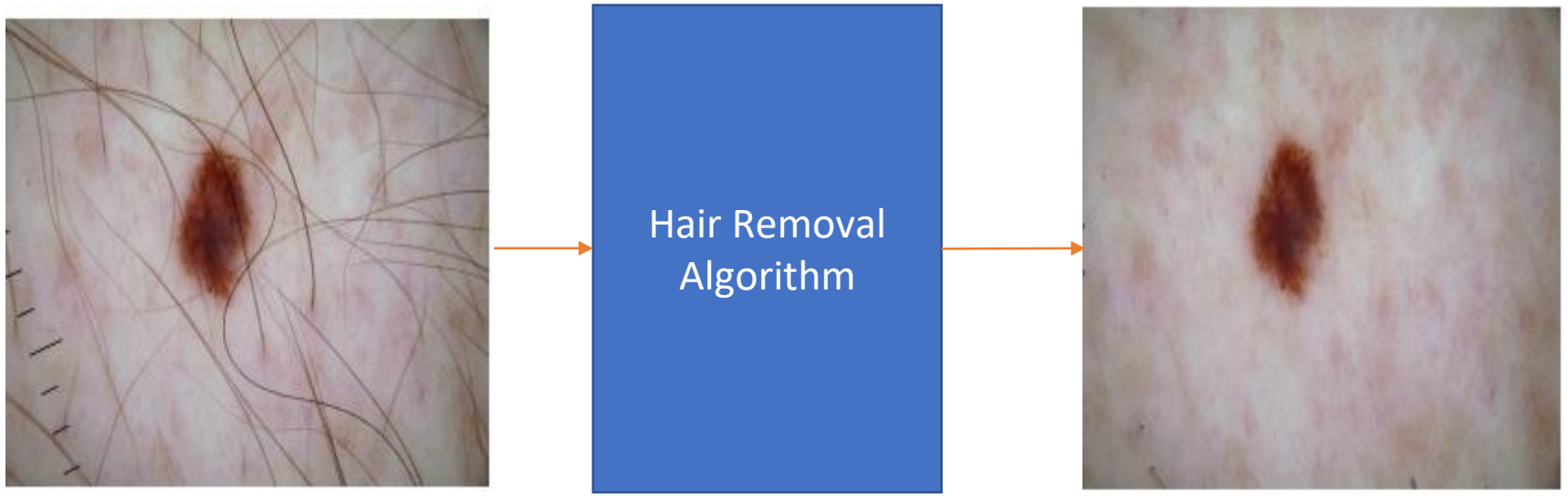


Image with Hair

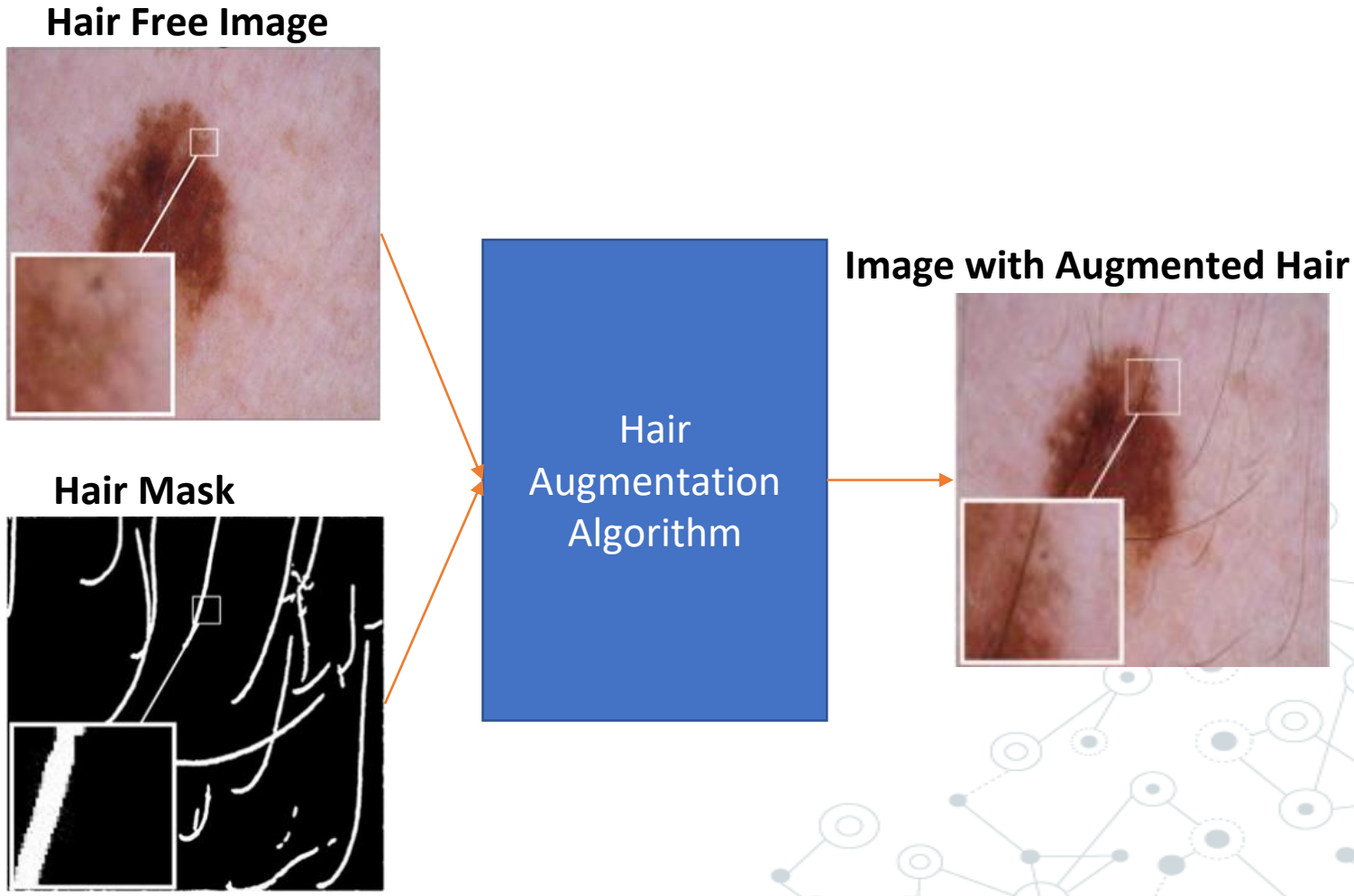
Hair Free Image

Disadvantage: will increase the computational complexity during real-time application

Research Plan
Skin hair augmentation

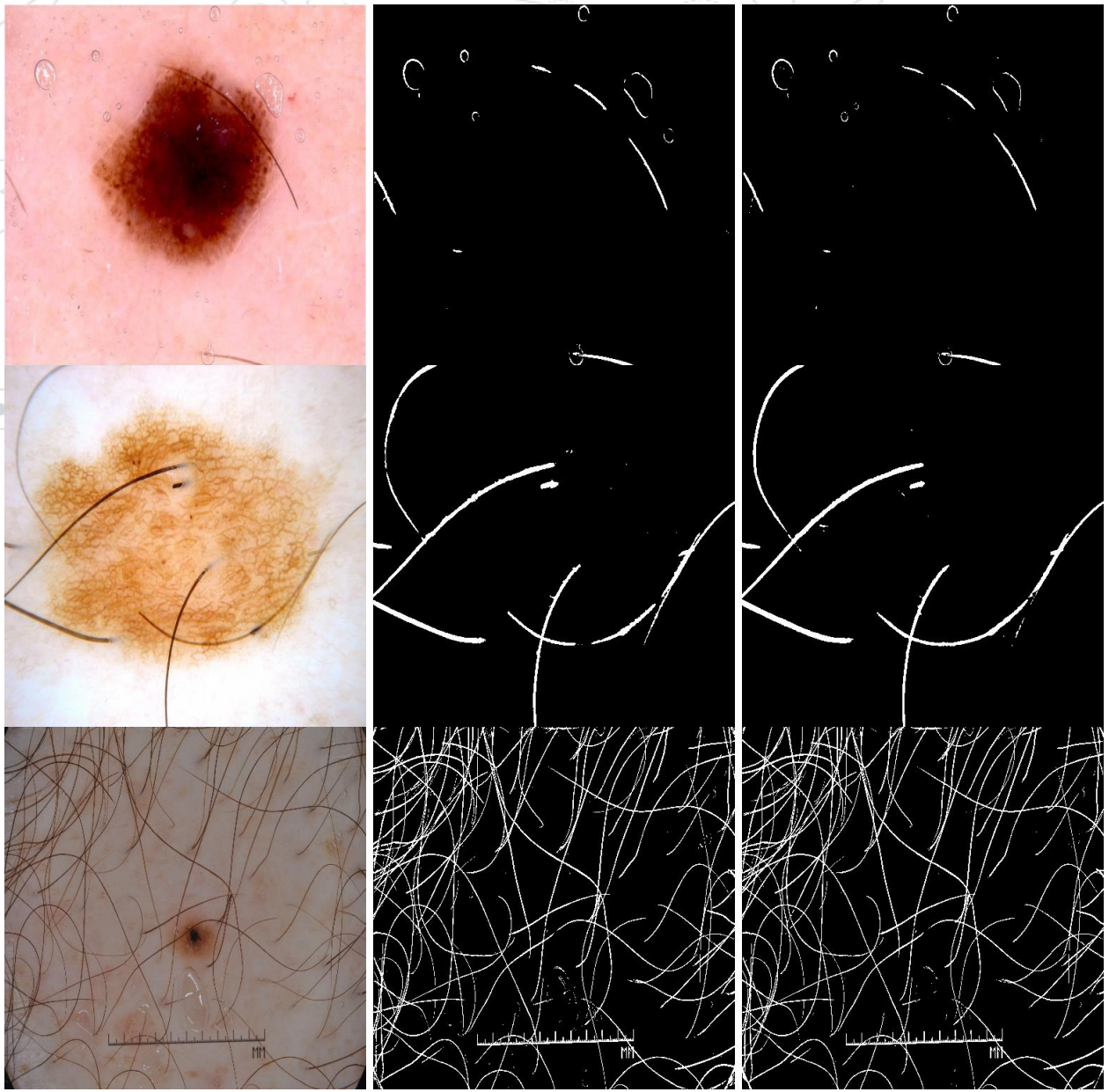
Skin Hair: Augmentation

- Mohamed et al. [5]



Research Scope: : Mask is created manually, with random curve or lines and segmentation. Generative Adversarial Network (GAN) can be utilized to automate the process.

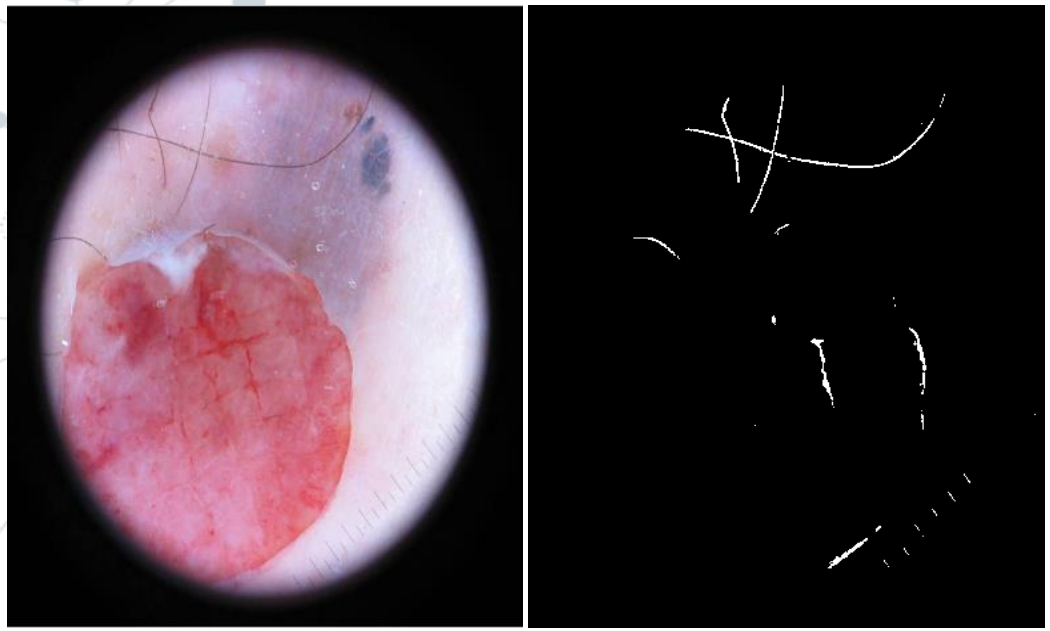
Research Plan
Skin hair augmentation



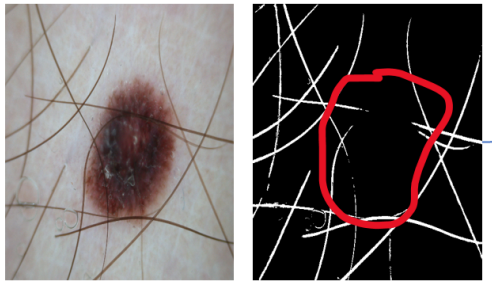
DATA:
-Curated after segmentation from images processing
-Claimed 306 images in paper but there were 288 unique (18 duplicate in train and test)
-Accuracy 0.984 (without augmentation and basic Unet)
-Ignored images with under segmentation

W. Li, A.N. Joseph Raj, T. Tjahjadi, Z. Zhuang, Digital hair removal by deep learning for skin lesion segmentation, Pattern Recognit. 117 (2021) 107994. https://doi.org/10.1016/J.PATCOG.2021.107994.

Research Plan
Skin hair augmentation



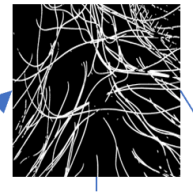
Hair mask dataset



Annotated mask dataset

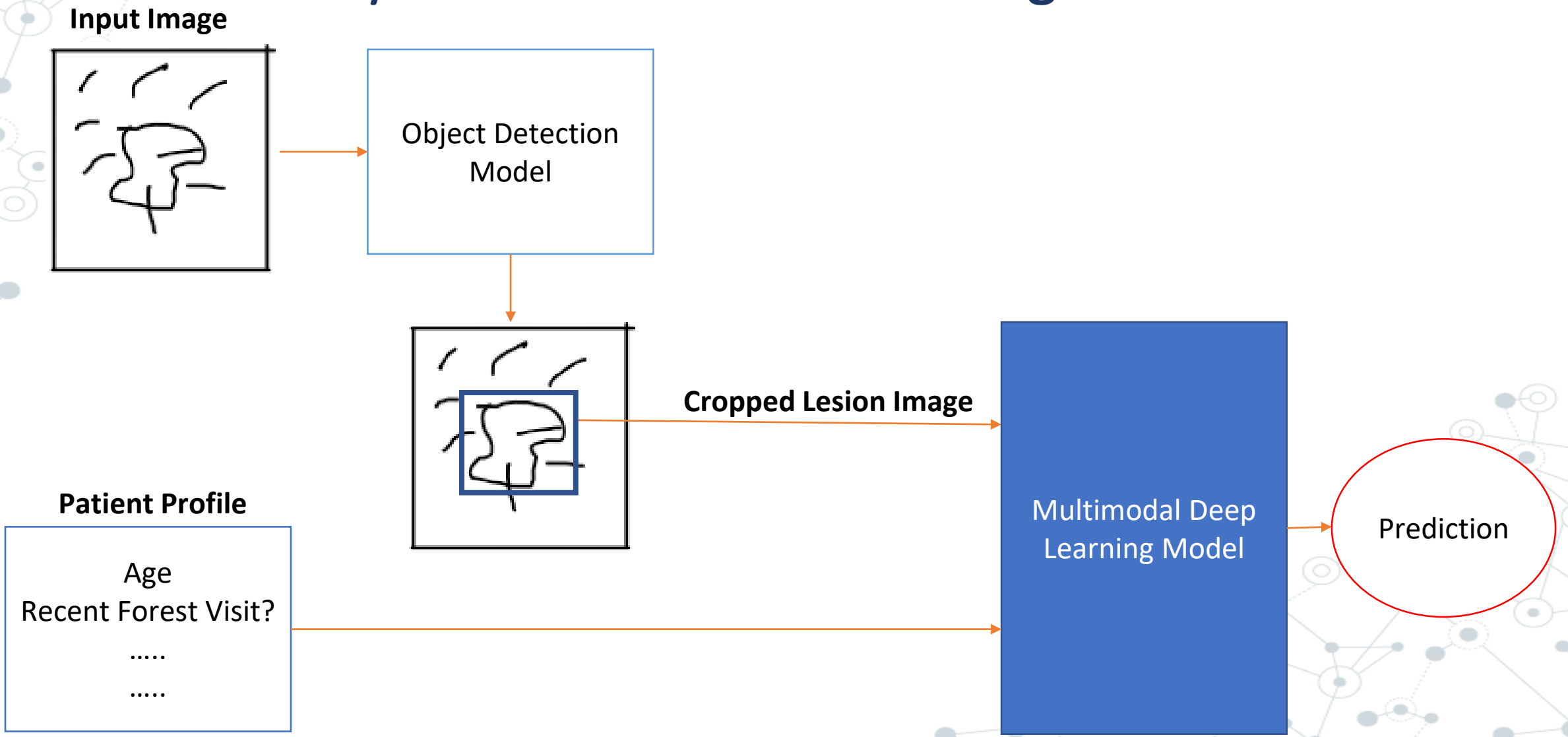
Manual correction
new dataset

Custom U-net
For segmentation



Generative Adversarial Network for realistic mask generation and filling up mask with realistic color

Research Plan
Planned system architecture with segmentation



Research Plan
Integrating Patient Profile

Questions	Réponses								Pondération	
		O Lesens	N Mrozek	J Beytout	C Theis	L Sauvat	M Vidal	C Cornille	Moyenne	std dev
Autres symptômes observés	Aucun	0	0	3	0	0	3	3	1.38	1.05
	Fièvre	-1	0	-1	1	1	3	0	0.75	0.90
	Fatigue	0	1	1	2	0	3	0	0.81	0.81
	Malaise	0	0	1	0	0	3	0	0.38	0.86
	Douleurs articulaires	1	1	-1	2	2	2	0	0.94	0.75
	Maux de tête	0	0	-1	0	2	1	0	0.56	0.79
	Démangeaison	-1	-1	-1	-1	1	-1	0	-0.34	0.76
Quelle est ou a été la taille maximale de cette tache	Moins de 1 cm	-1	-1	0	-1	-1	-1	-1	-0.69	0.58
	de 1 à 5 cm	1	1	1	0	1	1	1	1.00	0.50
	+ de 5 cm	3	2	2	2	3	3	3	2.44	0.61
	Je ne sais pas	0	0	0	0	0	0	0	0.00	0.35
La taille de cette tache augmente-t-elle ou a-t-elle augmentée progressivement	Oui	3	1	3	3	3	3	3	2.81	0.53
	Non	0	-1	-1	-1	-1	0	-1	-0.63	0.60
	Je ne sais pas	0	0	0	0	0	0	0	0.06	0.24
Avez-vous constaté une piqûre de tique à l'endroit de cette tache rouge au cours des 30 derniers jours ?	Oui	3	2	3	2	3	3	3	2.50	0.71
	Non	0	0	0	0	0	0	0	0.09	0.51
Fréquence des piqûres de tiques au cours des 30 derniers jours avant l'apparition de la tache rouge	Jamais	-1	-1	0	0	-1	0	0	-0.38	0.48
	1 fois	0	0	2	1	1	3	1	1.19	0.73
	2 à 5 fois	1	1	3	1	1	3	1	1.56	0.79
	plus de 5 fois	2	2	1	2	2	3	2	1.88	0.70
Activités de plein air au cours des 30 derniers jours avant l'apparition de la tache rouge	Oui	1	1	2	2	1	2	2	1.75	0.56
	Non	-1	-1	-1	-1	-1	0	0	-0.63	0.60

Research Plan
NAS : PSO with Selective Search

Selective Search

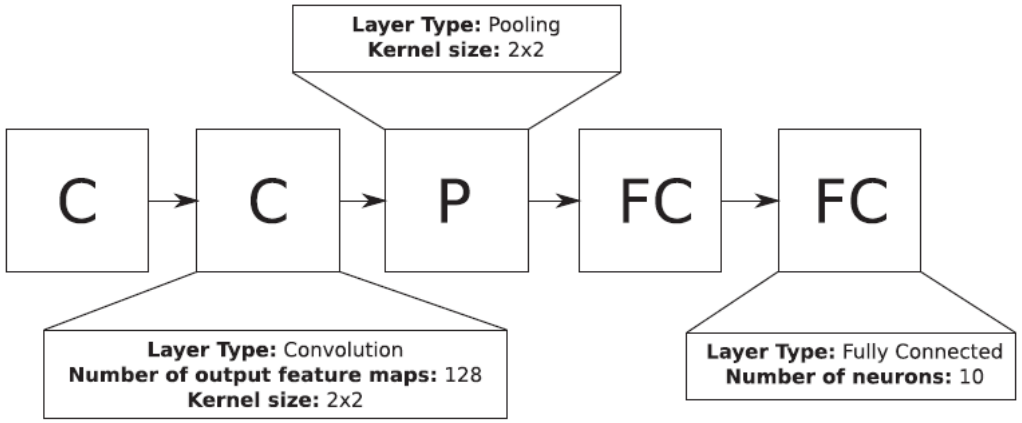
- Suppose, $SS = \{SO_1, SO_2, SO_3, \cdots, SO_n\}$ then the selective search can be written as

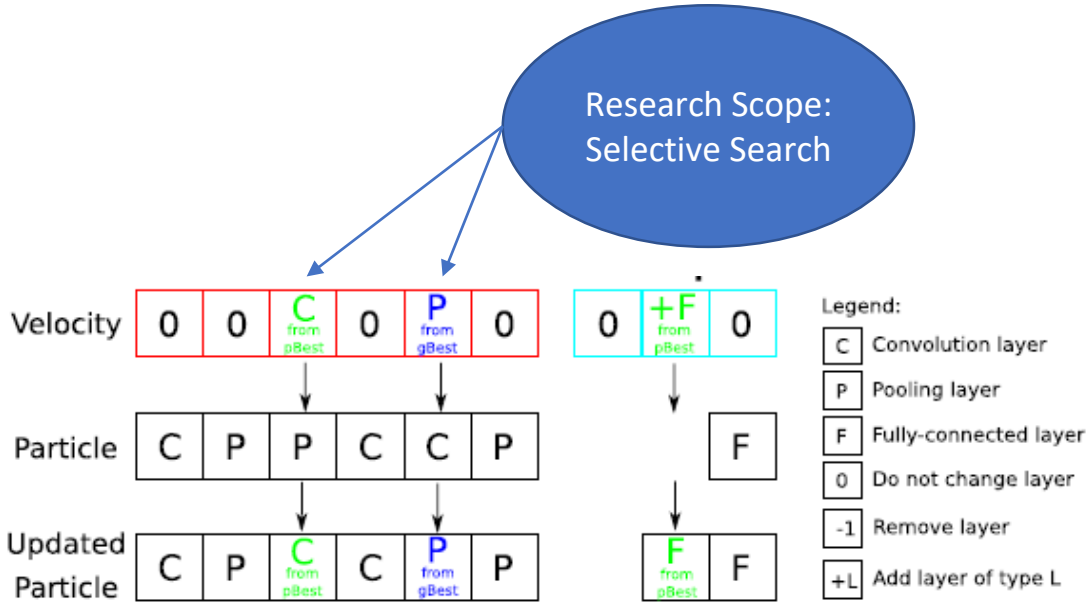
$$SC^1 = SC + SO_1$$
$$SC^2 = SC^1 + SO_2$$
$$\vdots$$
$$SC^n = SC^{n-1} + SO_n$$
$$SC = \max\{SC^f\}, \qquad f = 1,2,\cdots,n.$$

- Finally, the velocity is $SS = \{SO_1, SO_2, SO_3, \cdots, SO_f\}$, $1 \leq f \leq n$.

Hossain, S.I., et al., 2019. Optimization of University Course Scheduling Problem using Particle Swarm Optimization with Selective Search. Expert Syst. Appl. 127, 9–24. https://doi.org/10.1016/j.eswa.2019.02.026

Research Plan
NAS : PSO with Selective Search





Fernandes Junior, F.E., Yen, G.G., 2019. Particle swarm optimization of deep neural networks architectures for image classification. Swarm Evol. Comput. 49, 62–74. https://doi.org/10.1016/j.swevo.2019.05.010

References

[1] He, K., et al., 2016. Identity mappings in deep residual networks. LNCS. (including Subser. LNBI) 9908 LNCS, 630–645. https://doi.org/10.1007/978-3-319-46493-0_38

[2] Brinker, T.J., et al., 2019. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. Eur. J. Cancer 113, 47–54

[3] Russakovsky O. et al., 2015. ImageNet Large Scale Visual Recognition Challenge. Int. J. Comput. Vis. 115, 211–252. https://doi.org/10.1007/s11263-015-0816-y

[4] Tschandl, P., et al., 2018. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5. https://doi.org/10.1038/sdata.2018.161

[5] Attia, M., et al., 2020. Realistic hair simulator for skin lesion images: A novel benchemarking tool. Artif. Intell. Med. 108, 101933. https://doi.org/10.1016/j.artmed.2020.101933

Thank You