Deep Transfer Learning for Automated Diagnosis of Skin Lesions from Photographs

Doyoon Kim[1*], Emma Rocheteau[2*]
[1]Cleveland High School, California, US; doyoondeankim@gmail.com
[2]University of Cambridge, UK; ecr38@cam.ac.uk
[*]equal contribution

NEURAL INFORMATION PROCESSING SYSTEMS

Summary

- Melanoma is the deadliest form of skin cancer and requires expert knowledge by practitioners for diagnosis, which can be costly and inaccessible in certain parts of the world.
- With advancements in deep learning for improved diagnosis and increasing access to mobile technologies, this technology can be used for fast and efficient diagnosis.
- We compared the performance of several neural network architectures with and without transfer learning.

Model Architecture

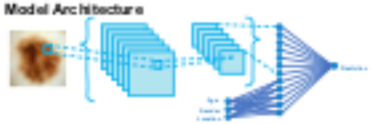


Figure 1: Model architecture. The CNN is different in each experiment.

- The CNN component (indicated in brackets) is different for each experiment. The static data is processed separately and concatenated to the CNN output before a final prediction is made.

Datasets

- We used the International Skin Imaging Collaboration (ISIC) 2020 dataset. It contains labelled photographs of skin lesions taken from various locations on the body.
- In total our data contained 37,648 skin lesion images.

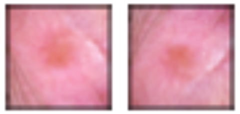
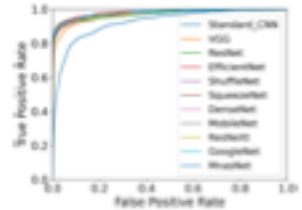


Figure 2: Example photographs in the training data. The original data is shown on the left and the augmented image is shown on the right.

- We performed data augmentation on the training data to introduce small variations in the form of random rotations, flipping, resizing, saturation shifts, etc. An example is shown above.

Accuracy Metrics



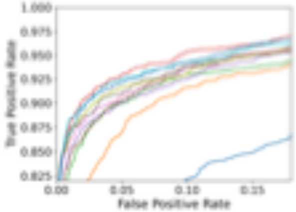


Figure 3: ROC curves of the CNN models and Standard CNN.

Integrated Gradients

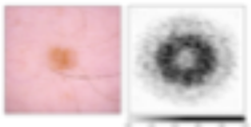


Figure 4: Skin image and its integrated gradient attribution (EfficientNet).

Results

	Model	Accuracy	AUROC	AUPRC	F1 Score
(a)	Standard CNN	[illegible]	[illegible]	[illegible]	[illegible]
	General Practitioners†	-	[illegible]	-	-
	VGG [30]	[illegible]	[illegible]	[illegible]	[illegible]
	SqueezeNet [10]	[illegible]	[illegible]	[illegible]	[illegible]
	ResNeXt [37]	[illegible]	[illegible]	[illegible]	[illegible]
	DenseNet [9]	[illegible]	[illegible]	[illegible]	[illegible]
	GoogleNet [33]	[illegible]	[illegible]	[illegible]	[illegible]
	ResNet-50 [6]	[illegible]	[illegible]	[illegible]	[illegible]
	MobileNet [28]	[illegible]	[illegible]	[illegible]	[illegible]
	MnasNet [35]	[illegible]	[illegible]	[illegible]	[illegible]
	ShuffleNet [38]	[illegible]	[illegible]	[illegible]	[illegible]
	EfficientNet [17]	[illegible]	[illegible]	[illegible]	[illegible]
	Dermatologists†	-	[illegible]	-	-
(b)	General Practitioners†	-	[illegible]	-	-
	VGG [30]	[illegible]	[illegible]	[illegible]	[illegible]
	ResNet-50 [6]	[illegible]	[illegible]	[illegible]	[illegible]
	ShuffleNet [38]	[illegible]	[illegible]	[illegible]	[illegible]
	SqueezeNet [10]	[illegible]	[illegible]	[illegible]	[illegible]
	DenseNet [9]	[illegible]	[illegible]	[illegible]	[illegible]
	Dermatologists†	-	[illegible]	-	-
	MobileNet [28]	[illegible]	[illegible]	[illegible]	[illegible]
	ResNeXt [37]	[illegible]	[illegible]	[illegible]	[illegible]
	GoogleNet [33]	[illegible]	[illegible]	[illegible]	[illegible]
	MnasNet [35]	[illegible]	[illegible]	[illegible]	[illegible]
	EfficientNet [17]	[illegible]	[illegible]	[illegible]	[illegible]

Table 1: Performance of the models. (a) without transfer learning (b) with.

- The error margins are 95 % confidence intervals (CIs). We report the accuracy, area under the receiver operating characteristic curve (AUROC), area under the precision recall curve (AUPRC) and the F1 Score.
- Within each table, the results are ordered from least to best performance.
- In table (b), if the result is statistically better than the model without transfer learning in a one tailed t-test ($p < 0.05$ * and $p < 0.001$ **), then it is indicated with stars.
- Results that significantly outperform general practitioners and dermatologists on AUROC (determined by a recent meta-analysis[24]) are indicated in green and blue respectively ($p < 0.05$).

Conclusion

- We have demonstrated the benefit of transfer learning for melanoma diagnosis. EfficientNet and MnasNet were capable of outperforming dermatologists.
- In future work, we would like to extend our binary classification task to multiclass (other skin lesions such as benign keratosis, basal cell carcinoma, etc.
- GitHub: https://github.com/aimadnus/Transfer_learning_melanoma