# Automated malignant melanoma detection using supervised contrastive learning



forward together sonke siya phambili saam vorentoe



Engineering | EyobuNjineli | Ingenieurswese

### Introduction

Stellenbosch UNIVERSITY IVUNIVESITHI UNIVERSITEIT

forward together sonke siya phambili saam vorentoe

This study investigates supervised contrastive learning and diverse encoder architectures for improving melanoma detection classification results.

- Challenges associated with melanoma detection
- Promise of computer-aided melanoma detection
- Favourable ImageNet classification results from 'Supervised Contrastive Learning' paper by Khosla et al. (2020).
- Encoder architectures explored:
  - Vision transformer
  - ResNet50
  - InceptionV3



# **Problem statement**





forward together sonke siya phambili saam vorentoe

- 1. Does supervised contrastive learning outperform traditional image classification models in detecting melanoma?
  - Hypothesis: Supervised contrastive learning has shown great promise in literature, producing state-of-the-art classification performance.
- 2. Do vision transformers yield superior classification performance over CNNs for the task of melanoma detection?
  - Hypothesis : Vision transformers have an enhanced long-range spatial awareness, resulting in impressive performance in recent literature.

# Baseline and Supervised Contrastive Learning Methodology



forward together sonke siya phambili saam vorentoe

#### Supervised Cross Entropy vs Self-supervised CL vs Supervised CL



#### Engineering | EyobuNjineli | Ingenieurswese

# Supervised contrastive learning & the SupCon Loss



**Supervised** contrastive loss pushes encodings from the same class closer together in the embedding space while pulling apart encodings from different classes.

Supervised contrastive loss function, SupCon:



# Supervised contrastive learning & the SupCon Loss



forward together sonke siya phambili saam vorentoe

**Supervised** contrastive loss pushes encodings from the same class closer together in the embedding space while pulling apart encodings from different classes.

#### Supervised contrastive loss function, SupCon:

Negate increasingly large values – small loss values, shrinking as representation space improves.



Increasingly large value divided by shrinking values > Increasingly large values

# Methodology

•

٠

•

•

Stellenbosch UNIVERSITY IYUNIVERSITHI UNIVERSITEIT

forward together sonke siya phambili saam vorentoe

**Class-weighted** Phase 1 batching scheme Load pretext task 25% class 'Melanoma' Data processing & weights Projection SupCon head loss augmentation Encoding Combine. 75% class 'Non-Melanoma' Training shuffle, Encoder Pretext training Phase 2 data rescale & augment **Staged training** Binary method cross-MLP entropy loss  $F_2$  prediction threshold



forward together sonke siya phambili saam vorentoe

#### Summary

#### 1. SupCon vs Baseline across encoders

- 2. Small vs larger batch size
- 3. Relative SupCon performance as batch size increases

Encoder Encoder parameters		VIT 34,276,288			<b>ResNet50</b> 23,564,800		INCEF 21,8	INCEPTIONV3 $21,802,784$	
TRAINING METHOD		BASE	SUPCON	В	ASE	SupCor	N BASE	SUPCON	
Batch size: 16	Threshold AUC Accuracy Recall <i>F</i> 2	$\begin{array}{c} 0.2455 \\ 0.8829 \\ 0.7109 \\ 0.9277 \\ 0.6210 \end{array}$	$\begin{array}{c} 0.2859 \\ 0.8759 \\ 0.7482 \\ 0.8735 \\ 0.6218 \end{array}$	0.3 0.8 0.7 0.8	$2131 \\ 8589 \\ 7728 \\ 8072 \\ 6052$	$\begin{array}{c} 0.2212 \\ 0.8889 \\ 0.7655 \\ 0.8916 \\ 0.6457 \end{array}$	$\begin{array}{c} 0.2212 \\ 0.8792 \\ 0.7568 \\ 0.8735 \\ 0.6288 \end{array}$	$0.2455 \\ 0.8740 \\ 0.6969 \\ 0.9398 \\ 0.6166$	
	Best score Mean score	$\begin{array}{c}2\\0.7856\end{array}$	$\begin{array}{c}2\\0.7798\end{array}$	0.'	1 7610	<b>3</b> 0.7979	$\frac{3}{0.7846}$	$\frac{1}{0.7818}$	
Batch size: 64	THRESHOLD AUC ACCURACY RECALL $F_2$	$\begin{array}{c} 0.2939 \\ 0.8646 \\ 0.7295 \\ 0.8675 \\ 0.6040 \end{array}$	$\begin{array}{c} 0.1646 \\ 0.7604 \\ 0.5396 \\ 0.9458 \\ 0.5223 \end{array}$	0.2 0.8 0.7 0.7	2455 8681 7688 7952 5951	$\begin{array}{c} 0.2374 \\ 0.8750 \\ 0.6969 \\ 0.9277 \\ 0.6106 \end{array}$	$\begin{array}{c} 0.1889 \\ 0.8703 \\ 0.7022 \\ 0.9036 \\ 0.6024 \end{array}$	$\begin{array}{c} 0.3101 \\ 0.8741 \\ 0.7722 \\ 0.8373 \\ 0.6216 \end{array}$	
	Best score Mean score	$\frac{3}{0.7664}$	$\begin{array}{c}1\\0.6920\end{array}$	0.'	1 7568	$\frac{3}{0.7776}$	$\frac{1}{0.7696}$	$\frac{3}{0.7763}$	



forward together sonke siya phambili saam vorentoe

#### Summary

#### SupCon vs Baseline across encoders

- 2. Small vs larger batch size
- 3. Relative SupCon performance as batch size increases

Encoder		VIT			ResNet50		INCEPTIONV3		
Encoder parameters		$34,\!276,\!288$			$23,\!564,\!800$		21	$21,\!802,\!784$	
TRAINING METHOD		BASE	SupCon	В	BASE	SUPCON	BASE	e Sup	Con
BATCH SIZE:	Threshold	0.2455	0.2859	0.	2131	0.2212	0.221	2 0.2	455
16	AUC	0.8829	0.8759	0.	8589	0.8889	0.879	0.8	740
	ACCURACY	0.7109	0.7482	0.	7728	0.7655	0.756	0.6	969
	Recall	0.9277	0.8735	0.	8072	0.8916	0.873	0.9	398
	$F_2$	0.6210	0.6218	0.	6052	0.6457	0.628	.6 0.6	166
	Best score	2	2		1	3	3	-	1
	MEAN SCORE	0.7856	0.7798	0.	7610	0.7979	0.784	6 0.7	818
BATCH SIZE:	Threshold	0.2939	0.1646	0.	2455	0.2374	0.188	.3 0.3	101
64	AUC	0.8646	0.7604	0.	8681	0.8750	0.870	0.8	741
	ACCURACY	0.7295	0.5396	0.	7688	0.6969	0.702	22 - 0.7	722
	RECALL	0.8675	0.9458	0.	7952	0.9277	0.903	0.8	373
	$F_2$	0.6040	0.5223	0.	5951	0.6106	0.602	24 0.6	216
	Best score	3	1		1	3	1	ę	3
	MEAN SCORE	0.7664	0.6920	0.	7568	0.7776	0.769	06 0.7	763



forward together sonke siya phambili saam vorentoe

#### Summary

#### 1. SupCon vs Baseline across encoders

- 2. Small vs larger batch size
- 3. Relative SupCon performance as batch size increases

Encoder Encoder parameters		VIT 34.276.288		$\begin{array}{c} \text{ResNet50} \\ \text{23.564.800} \end{array}$		INCEPTIONV3 21.802.784	
TRAINING METHOD		BASE	SUPCON	BASE	SUPCON	BASE	SUPCON
Batch size: 16	Threshold AUC Accuracy Recall E-	$\begin{array}{c} 0.2455 \\ 0.8829 \\ 0.7109 \\ 0.9277 \\ 0.6210 \end{array}$	$\begin{array}{c} 0.2859 \\ 0.8759 \\ 0.7482 \\ 0.8735 \\ 0.6218 \end{array}$	$\begin{array}{c} 0.2131 \\ 0.8589 \\ 0.7728 \\ 0.8072 \\ 0.6052 \end{array}$	$\begin{array}{c} 0.2212 \\ 0.8889 \\ 0.7655 \\ 0.8916 \\ 0.6457 \end{array}$	$\begin{array}{c} 0.2212 \\ 0.8792 \\ 0.7568 \\ 0.8735 \\ 0.6288 \end{array}$	$\begin{array}{c} 0.2455 \\ 0.8740 \\ 0.6969 \\ 0.9398 \\ 0.6166 \end{array}$
	Best score Mean score	2 0.7856	2 0.7798	1 0.7610	3 0.7979	3 0.7846	1 0.7818
Batch size: 64	THRESHOLD AUC ACCURACY RECALL $F_2$	$\begin{array}{c} 0.2939 \\ 0.8646 \\ 0.7295 \\ 0.8675 \\ 0.6040 \end{array}$	$\begin{array}{c} 0.1646 \\ 0.7604 \\ 0.5396 \\ 0.9458 \\ 0.5223 \end{array}$	$\begin{array}{c} 0.2455 \\ 0.8681 \\ 0.7688 \\ 0.7952 \\ 0.5951 \end{array}$	$\begin{array}{c} 0.2374 \\ 0.8750 \\ 0.6969 \\ 0.9277 \\ 0.6106 \end{array}$	$\begin{array}{c} 0.1889 \\ 0.8703 \\ 0.7022 \\ 0.9036 \\ 0.6024 \end{array}$	$\begin{array}{c} 0.3101 \\ 0.8741 \\ 0.7722 \\ 0.8373 \\ 0.6216 \end{array}$
	Best score Mean score	$\frac{3}{0.7664}$	$\frac{1}{0.6920}$	$\frac{1}{0.7568}$	$\frac{3}{0.7776}$	$\frac{1}{0.7696}$	$\frac{3}{0.7763}$



forward together sonke siya phambili saam vorentoe

#### Summary

#### 1. SupCon vs Baseline across encoders

- 2. Small vs larger batch size
- 3. Relative SupCon performance as batch size increases

Encoder		VIT		Res	Net50	INCEPTIONV3		
Encoder parameters		$34,\!276,\!288$		$23,\!564,\!800$		$21,\!802,\!784$		
TRAINING METHOD		BASE	SUPCON	BASE	SupCon	BASE	SupCon	
BATCH SIZE:	Threshold	0.2455	0.2859	0.2131	0.2212	0.2212	0.2455	
16	AUC	0.8829	0.8759	0.8589	0.8889	0.8792	0.8740	
	ACCURACY	0.7109	0.7482	0.7728	0.7655	0.7568	0.6969	
	RECALL	0.9277	0.8735	0.8072	0.8916	0.8735	0.9398	
	$F_2$	0.6210	0.6218	0.6052	0.6457	0.6288	0.6166	
	Best score	2	2	1	3	3	1	
	Mean score	0.7856	0.7798	0.7610	0.7979	0.7846	0.7818	
BATCH SIZE:	Threshold	0.2939	0.1646	0.2455	0.2374	0.1889	0.3101	
64	AUC	0.8646	0.7604	0.8681	0.8750	0.8703	0.8741	
	Accuracy	0.7295	0.5396	0.7688	0.6969	0.7022	0.7722	
	RECALL	0.8675	0.9458	0.7952	0.9277	0.9036	0.8373	
	$F_2$	0.6040	0.5223	0.5951	0.6106	0.6024	0.6216	
	Best score	3	1	1	3	1	3	
	Mean score	0.7664	0.6920	0.7568	0.7776	0.7696	0.7763	

#### **Test Results**



- AUC: 0.8569
- Accuracy: 0.7055
- Recall: 0.8304
- F2: 0.5717
- Mean score: 0.7411



# **Problem statement**





forward together sonke siya phambili saam vorentoe

- Does supervised contrastive learning outperform traditional image classification models in detecting melanoma?
  - Finding: Supervised contrastive learning does not consistently produce superior melanoma detection performance.
- 2. Do vision transformers yield superior classification performance over CNNs for the task of melanoma detection?
  - Finding : ViT encoders do not necessarily outperform CNN architectures.



forward together sonke siya phambili saam vorentoe

