Skin pathology characterization by infrared thermography.

by Jan Verstockt*, Warre Clarys*, Mathias Bonmarin**, Filip Thiessen***, Lieve Brochez****, Gunther Steenackers*

* *University of Antwerp, InViLab research group, Groenenborgerlaan 171, B-2020 Antwerp, Belgium*
** *School of Engineering, Zurich University of Applied Sciences (ZHAW)*
*** *Gynecological Oncology Unit, Department of Obstetrics and Gynecology, Multidisciplinary Breast Clinic, Antwerp University Hospital, University of Antwerp, Antwerp, Belgium.*
**** *Department of Dermatology, Ghent University Hospital, Belgium; Dermatology Research Unit, Ghent, Belgium; Department of Medical Oncology, Ghent University Hospital, Belgium; Immuno-Oncology Network Ghent (ION Ghent), Belgium*

Abstract

Active infrared thermography (IRT) was investigated as a non-invasive method for localising and characterising cutaneous lesions through transient thermal responses following controlled cooling. A five-layer finite-element skin model based on Pennes' bioheat equation generated synthetic datasets for lesion parameter estimation using machine learning. CNN, U-Net and 3D-CNN architectures achieved sub-millimetric accuracy under simulated conditions, while experimental validation with skin phantoms demonstrated reproducible cooling and reheating measurements. The HypIRskin prototype, developed with Zurich University of Applied Sciences, enabled standardised acquisition within approximately three minutes. Results indicate that active IRT can provide diagnostically relevant information, warranting further clinical validation.

1. Introduction

Skin cancer represents one of the fastest-growing cancer types worldwide [1]. According to recent epidemiological studies, incidence rates continue to increase throughout Europe, mainly due to cumulative ultraviolet exposure and ageing populations [2]. Among all skin malignancies, melanoma remains the most aggressive form because of its metastatic potential and associated mortality [3]. Current diagnostic pathways rely primarily on visual inspection, dermoscopy and histopathological confirmation through biopsy [4, 5]. Although dermoscopy significantly improves diagnostic performance, the specificity remains limited, resulting in unnecessary excisions and increased healthcare costs [6].

The accurate detection and characterization of cutaneous lesions remain a significant challenge in contemporary dermatology. While visual inspection and dermoscopy are the current clinical standards, they are inherently subjective and often fail to provide quantitative information regarding the subsurface geometry of a lesion, such as its thickness and depth of invasion. [7] Given that parameters like the Breslow thickness are critical prognostic indicators for malignancies such as melanoma, there is an urgent need for non-invasive, objective imaging modalities capable of depth-resolved skin analysis [8, 9].

Active Infrared Thermography (IRT) has emerged as a promising candidate to address these limitations [10]. Unlike passive thermography, which merely records the steady-state surface temperature of the skin, active IRT involves the application of an external thermal stimulus, typically controlled cooling, to induce a transient thermal response. Because the rate of reheating is governed by the underlying tissue's thermal properties (such as conductivity and diffusivity) and physiological factors (such as blood perfusion), the resulting spatiotemporal temperature profiles contain encoded information about the lesion's internal structure [11]. Dynamic Infrared Thermography (DIRT) has attracted growing interest as a non-invasive imaging technique capable of detecting physiological alterations associated with malignant tissue. Increased vascularization, angiogenesis and metabolic activity modify local heat transfer characteristics and may therefore produce measurable differences in thermal recovery following external thermal excitation [12].

However, the clinical translation of active IRT is often hindered by the complexity of the inverse problem: reconstructing three-dimensional lesion parameters from two-dimensional surface measurements. Traditional analytical methods frequently struggle with the heterogeneous and multi-layered nature of human skin. To overcome this, physics-based numerical modeling can be combined with deep learning architectures [13]. By employing a five-layer parametric finite-element model based on Pennes' bioheat equation, it is possible to generate extensive synthetic datasets that account for biological variability [14]. These datasets provide the foundation for training advanced neural networks, including Convolutional Neural Networks (CNNs) and U-Net architectures, to perform automated localization and parameter estimation. Despite encouraging results reported in literature, clinical adoption remains limited due to the absence of standardized measurement procedures and dedicated instrumentation. Variability in cooling methods, camera positioning and acquisition protocols negatively affects measurement reproducibility. To address these limitations, the HypIRskin project developed

License: https://creativecommons.org/licenses/by-nd/4.0/deed.en

a dedicated clinical DIRT platform specifically designed for skin cancer assessment.

This paper presents an end-to-end investigation into the feasibility of characterizing skin pathology through an integrated active IRT framework. We detail the development of a multi-camera imaging prototype, the "HypIRskin," and evaluate the performance of various deep learning models in recovering lesion diameter, depth, and shape. Through experimental validation using skin-mimicking phantoms and simulated data, we demonstrate that active IRT, supported by physics-informed machine learning, offers a scalable and robust pathway for objective skin cancer screening and surgical planning.

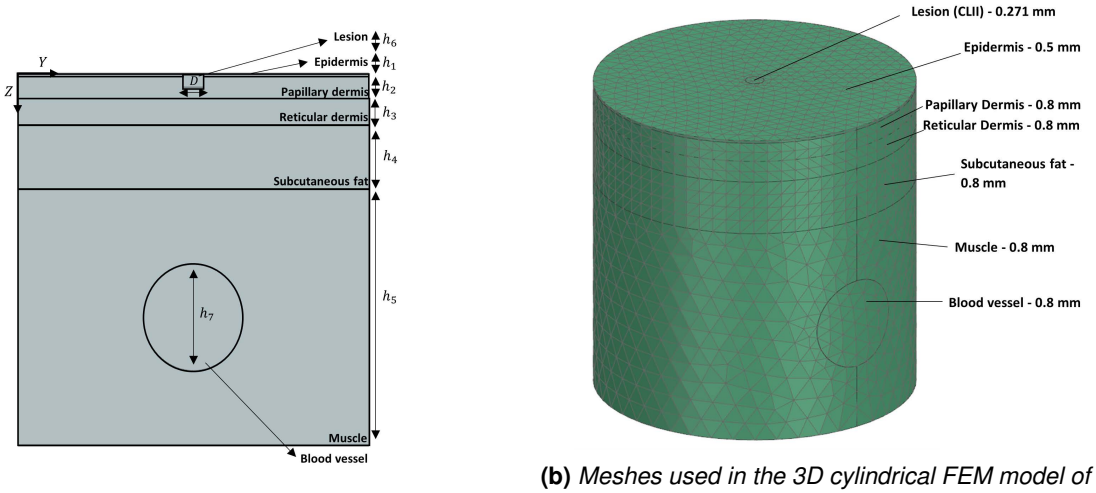
2. Materials and methods

2.1 Finite Element Modeling of Human Skin for Infrared Thermography-Based Melanoma Detection

2.1.1 Model geometry and layer architecture

A three-dimensional (3D) cylindrical finite element model of human skin was developed using Siemens Simcenter 3D. A cylindrical geometry was preferred over two-dimensional or cubic three-dimensional alternatives, as prior work has demonstrated that 2D models tend to overestimate temperature differences [15], while cylindrical 3D models reduce computational cost by imposing a single lateral boundary condition rather than four [15]. The diameter of the skin cylinder was set to 30 mm, corresponding to the field of view of the Optris Xi400 thermal camera with microscope optics at an imaging distance of 100 mm.

The model comprises five skin layers (Figure 1): the epidermis, the papillary dermis, the reticular dermis, the subcutaneous fat, and the muscle layer. The stratum corneum and bone layer were intentionally excluded; the former because its 0.01 mm thickness renders its thermal influence negligible, and the latter due to the absence of thermally relevant properties such as blood perfusion and metabolic heat generation. Within the top layer, a centrally positioned cylindrical lesion is embedded, with adjustable diameter (1–15 mm) and depth (0.05–2.65 mm), covering the range of clinically relevant melanoma stages defined by the Clark staging system.



(a) *Layers of the skin transferred to a 3D skin model*

(b) *Meshes used in the 3D cylindrical FEM model of the skin*

Fig. 1. *Finite Element skin model with 5 layers as modelled in Siemens Simcenter 3D*

Thermophysical parameters for each layer, including specific heat capacity c, thermal conductivity k, density ρ, blood perfusion rate ω_b, and metabolic heat generation Q_m, were adopted from the validated reference values established by Çetingül and Herman (Table 1) [16]. Tumour parameters reflect the elevated metabolic activity of malignant tissue: blood perfusion is five times higher and metabolic heat generation ten times higher than in surrounding healthy tissue, consistent with experimental findings reported by Song et al. [17] and theoretical assumptions by Agrawal [18].

All model parameters were defined as expressions in Siemens Simcenter 3D to allow external modification during automated simulation runs, enabling systematic variation across both the parameter sensitivity study and the tumour study.

Table 1. *Thermophysical parameters of the five-layer skin model, adopted from Çetingül and Herman.*

Layer	h (mm)	c (J/kg·K)	k (W/m·K)	ρ (kg/m^3)	ω_b (1/s)	Q_m (W/m^3)
Epidermis	0.1	3589	0.235	1200	0	0
Papillary dermis	0.7	3300	0.445	1200	0.0002	368.1
Reticular dermis	0.8	3300	0.445	1200	0.0013	368.1
Subcutaneous fat	2.0	2674	0.185	1000	0.0001	368.3
Muscle	8.0	3800	0.510	1085	0.0027	684.2
Lesion	0.2–1.0	3852	0.558	1030	0.0063	3680

2.1.2 Governing Equation: Pennes' Bioheat Equation

Heat transfer within the skin model is governed by Pennes' bioheat equation [19], which describes the transient temperature distribution in living soft tissue:

$$\rho C \frac{\partial T}{\partial t} = k\nabla^2 T + \omega_b C_b (T_a - T) + Q_m + Q_r(x, y, z, t) \tag{1}$$

where T is tissue temperature, ρ tissue density, C specific heat capacity, k thermal conductivity, ω_b blood perfusion rate, C_b specific heat of blood, T_a arterial temperature, Q_m metabolic heat generation, and Q_r any externally applied regional heat source. The equation is applied independently to each layer in all three spatial dimensions. Two boundary conditions are imposed: (i) temperature continuity at inter-layer interfaces, ensuring conservation of heat flux across layer boundaries; and (ii) a zero-flux lateral boundary condition at the outer cylindrical wall, reflecting the assumption that the surrounding tissue is at a uniform, constant temperature. At the bottom face of the model, the core body temperature is held constant at 37 ℃; at the top surface, convective heat exchange with ambient air is modeled using a convective heat transfer coefficient of 10 $\frac{W}{m^2 K}$ and a skin emissivity of 0.98.

2.1.3 Three-phase simulation protocol

Each simulation consists of three sequential phases as shown in Figure 2:

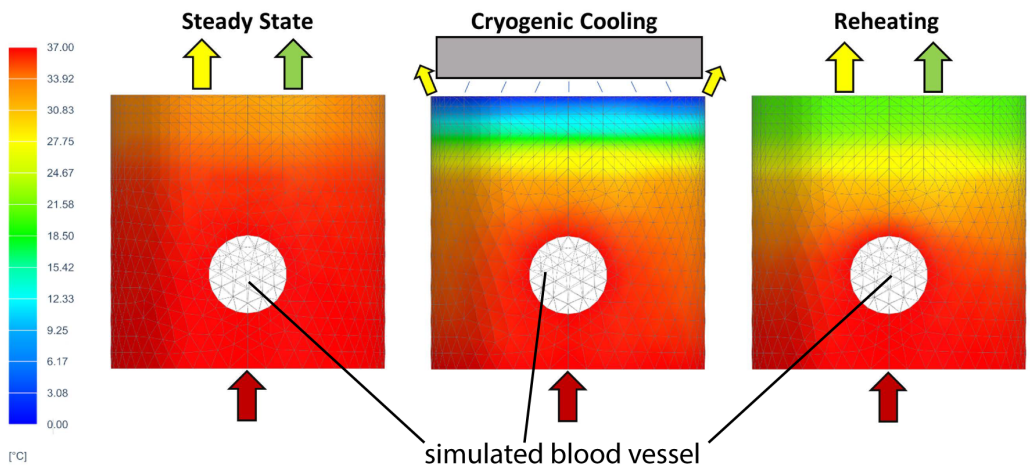


Fig. 2. *The model consists of three phases: steady-state phase, cooling phase, and natural reheating phase. The yellow arrow represents the emissivity of the skin surface, the green arrow indicates convection to the ambient air, and the red arrow denotes the core body temperature. The white circle depicts a simulated blood vessel, which may be included or excluded in the simulations.*

Steady state A non-transient simulation establishing the thermal equilibrium of the skin under resting conditions (ambient temperature 21 ℃, body core temperature 37 ℃). This phase serves as the initial condition for subsequent dynamic phases.

Cooling A cryogenic cooling phase using the simulated Zimmer Cryo 6 device. Three intensity settings (levels 3, 6, and 9) were evaluated on a healthy skin model to identify the setting that achieves sufficient cooling depth, defined as a 5 °C temperature drop at 2.65 mm depth, while maintaining the most tolerable surface temperature. Level 3 (airflow: 1000 L/min at -30 °C; surface velocity: 23.58 m/s) was selected for simulation, resulting in surface cooling to approximately 13 °C over 39 seconds (40 simulation iterations at 1 s/iteration).

Reheating Following the cooling, the skin model is re-exposed to the same ambient conditions as in the steady-state phase. Temperature differences between tumour and healthy tissue are most diagnostically informative during this recovery phase. The reheating phase is simulated for 210 seconds (211 iterations at 1 s/iteration), concluding when the surface temperature returns to approximately 30 °C, at which point the model approaches thermal equilibrium and further data becomes less informative. The complete simulation produces 252 time frames per tumour variant.

2.1.4 Parameter Sensitivity study

A one-at-a-time parameter sensitivity analysis was performed to identify which physical parameters most strongly influence the simulated surface temperature. For each of the 25 parameters, five physical properties across five skin layers, two simulations were run using the lower and upper bounds of clinically realistic variation intervals (Table 1). The output metric was the mean temperature deviation, averaged across the steady-state and reheating phases and across two reference surface points: the tumour centre and a point at five tumour radii, the latter serving as a healthy tissue reference. Results confirm that skin layer thickness h and blood perfusion rate ω_b exert the greatest influence on surface temperature, consistent with earlier findings by Çetingül and Herman [20]. The subcutaneous fat layer has the largest individual contribution, attributable to its exceptionally low thermal conductivity. Metabolic heat generation Q_m has the least impact across all layers, as its contribution is small relative to the convective heat transfer associated with blood perfusion. The close agreement between the present sensitivity results and those of Çetingül et al. validates the implemented skin model. Simulation management and parameter variation were automated using Siemens HEEDS, which controlled Siemens Simcenter 3D and extracted output files without manual intervention. Post-processing was performed in MATLAB.

2.2 HypIRskin prototype

Despite growing evidence that dynamic infrared thermography (DIRT) can reveal diagnostically relevant characteristics of skin lesions, no standardised, quantitative, and non-invasive IRT-based method has yet been established for routine clinical use. Prior research has employed a variety of cooling excitation approaches, including cold air from vortex tubes, fans, and air-conditioning units, as well as conductive methods using cooled gel packs or aluminium plates, and cameras have typically been tripod-mounted or handheld, without systematic control of imaging distance, angle, or cooling uniformity. Reproducibility has been further complicated by the need for visible markers to delimit the region of interest and compensate for patient movement during data acquisition. To address these limitations, HypIRskin (Figure 3) was developed as a self-contained, integrated measurement platform that combines a thermal camera, a controlled cryogenic cooling unit, an RGB camera, laser alignment, environmental sensing, and dedicated acquisition software within a single measurement head [21]. The design prioritises three properties essential for clinical translation: measurement reproducibility, ease of use in a dermatological or surgical setting, and a fully automated acquisition protocol that minimises operator-dependent variability.

2.2.1 Hardware components

The HypIRskin device is built around a mini-PC (Udoo Bolt V8, AMD Ryzen Embedded V1605B quad-core processor) that centrally orchestrates all hardware and software components. The hardware sub-systems are integrated into a compact measurement head (Figure 2) and a main device body connected via USB and Ethernet.

Thermal camera The primary imaging sensor is an Optris Xi400 long-wave infrared camera (Optris GmbH, Berlin, Germany) fitted with microscope optics (angular field of view: 18° × 14°). The camera features an uncooled microbolometer detector with a 17 µm pixel pitch and a thermal sensitivity (NETD) of <80 mK, capturing 16-bit RAW thermal sequences at 382 × 288 pixels and 27 Hz. A distance holder integrated into the measurement head maintains a fixed standoff of 100 mm from the skin surface, yielding a field of view of 34.2 × 25.7 mm. This fixed geometry directly determines the 30 mm diameter used in the FEM skin model described in Section 2.1. The microscope optics enable sub-millimetre features of skin lesions to be resolved at high contrast.

Cooling unit Cryogenic cooling is provided by a Zimmer Cryo C cold-air device, a CE-marked Class IIa medical device under Medical Device Directive Annex IX, delivering filtered, oil-free, moisture-free air at -30 °C with a flow rate adjustable across nine levels up to 1000 L/min. Air temperature at the nozzle outlet is actively monitored via a Platinum Thin Film Pt100 RTD sensor (Adafruit MAX31865 amplifier, precision $\geq$ 0.5 °C) and regulated by a PID controller implemented in Python, which drives a 690 W heating coil via a Eurotherm TE10a thyristor controller and a National Instruments DAC. This closed-loop temperature control enables the air outlet temperature to be varied from -30 °C to +60 °C, allowing not

Fig. 3. *left: CAD model of the device, middle: CAD model of the internal structure of the device, right: Prototype device as built*

only standard cooling protocols but also lock-in thermography sequences. During prototype testing, level 5 of the Cryo C was used for in-vivo measurements, producing a temperature differential of approximately 0.8 °C between malignant and surrounding healthy tissue during reheating.

RGB camera and motion compensation A Daheng Imaging PoE camera (MER-133-54GC-P, 1280 × 960 pixels, 54 fps) with a polarising filter is synchronised with the thermal camera to capture simultaneous RGB frames. The RGB stream serves two functions: real-time monitoring of the skin lesion and post-acquisition motion compensation via tracking algorithms, eliminating the need for a fiducial marker visible in both modalities.

Laser alignment system Two red laser diodes (650 nm, 7 mW each) are angled so that their beams converge to form a cross at the centre of the measurement area when the device is correctly positioned at 100 mm from the skin. This facilitates rapid and repeatable alignment over the lesion without contact.

Illumination and environmental sensing LED arrays with polarising filters provide controlled illumination of the measurement area while minimising specular reflections from the skin surface. A Bosch BME280 sensor monitoring ambient temperature, relative humidity, and barometric pressure is positioned close to the nozzle outlet to characterise airflow conditions. All peripheral components within the measurement head are governed by a custom-designed PCB incorporating a MAX31865 RTD-to-digital converter and an Arduino Nano (ATmega328), which communicates with the mini-PC via serial protocol and independently manages button states, LED/laser toggling, and temperature acquisition.

2.2.2 Camera Calibration and Image Registration

Accurate spatial correspondence between the RGB and thermal images is required for motion compensation and for overlaying visible lesion morphology with thermographic data. Standard checkerboard calibration targets are unsuitable for thermal cameras due to insufficient contrast between black and white squares in the long-wave infrared band. A heated UV-printed Dibond target was therefore developed, providing adequate thermal contrast for corner detection in both imaging modalities. Geometric (intrinsic) calibration of each camera was performed independently, yielding lens distortion maps and intrinsic parameter matrices. After undistortion, a planar homography was computed from corresponding checkerboard corner locations in the rectified RGB and thermal images. The homography matrix H maps pixel coordinates from the RGB image plane to the thermal image plane. Once registered, the RGB image, converted to greyscale and encoded in the red channel, and the thermal image, encoded in the blue channel, are overlaid in a composite image (Figure 4), enabling direct visual comparison of lesion morphology and thermal signature.

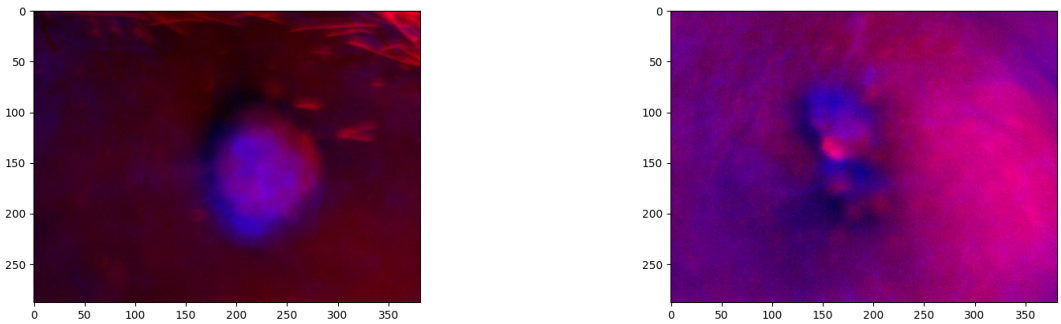


Fig. 4. *Melanoma lesions captured with HypIRskin device. The red layer depicts the rgb layer which is transformed to gray values and colored red. The blue color shows the overlay of the thermal data onto the rgb layer.*

2.3 Deep learning: Diameter, Depth, and Lesion Segmentation from Simulated Reheating Sequences

Surgical excision of cutaneous melanoma follows margin guidelines based on Breslow thickness: 5 mm for melanoma in situ, 10 mm for lesions below 1.0 mm, 10–20 mm for 1.0–2.0 mm lesions, and 20 mm for thicker lesions. These conservative margins reduce the risk of incomplete excision but contribute to unnecessary tissue removal and scarring. A pre-surgical method quantifying lesion diameter and depth with sub-millimetre accuracy could allow margin planning to be tailored to actual tumour geometry. Dynamic infrared thermography (DIRT) detects subsurface thermal anomalies produced by the elevated blood perfusion and metabolic heat generation of malignant tissue, which manifest as a measurable surface temperature differential during skin reheating after cryogenic cooling. The spatiotemporal structure of this differential encodes information about lesion depth, diameter, and areal extent. Deep learning offers a principled means of learning this mapping directly from thermographic image sequences. This section describes the development and evaluation of three complementary models trained on FEM-simulated reheating data: a CNN regression model predicting tumour diameter and depth; a U-Net segmentation model estimating tumour position and area; and a 3D CNN performing end-to-end spatiotemporal analysis. Performance is benchmarked against the clinically applied surgical excision margins.

2.3.1 Dataset

The dataset comprises 270 HDF5 files derived from the FEM simulation study (Section 2), each containing the 211-frame reheating sequence of a single lesion geometry at 1000 × 1000 pixel resolution. Tumour diameters range from 1–10 mm in 1 mm increments and depths from 0.05–2.65 mm in 0.10 mm increments, covering the full Clark staging range. Images were cropped to 720 × 720 pixels to remove peripheral NaN artefacts. For the CNN regression model, Principal Component Analysis (PCA) was applied across the 211 frames, with three principal components selected as the optimal balance between explained variance and memory footprint. For the U-Net, five frames at indices 0, 50, 100, 150, and 200 were sampled to preserve temporal progression without dimensionality reduction, and normalised per-sequence to [0, 1]. Binary ground-truth segmentation masks were generated from tumour centre coordinates, with positions tracked across spatial crops.

Data augmentation Random cropping, Gaussian blur, and binary noise was applied to all three splits (80% training / 10% validation / 10% test), rather than the training set alone. This reflects the idealised nature of the synthetic data and ensures that evaluation accounts for realistic imaging variability.

2.3.2 CNN Regression model

The regression CNN takes 3-channel PCA images (720 × 720 pixels) as input and predicts tumour diameter and depth simultaneously through two parallel output heads. After a sequence of convolutional, max-pooling, and ReLU layers, the feature vector is reduced from 291,600 to 512 dimensions by a fully connected layer (FC1), with two separate FC2 heads producing the two predictions. The training loss is the weighted sum of the MSE losses for both outputs. Hyperparameters were tuned by sequential grid search (five independent runs per configuration) on an NVIDIA RTX 4070 GPU. The final configuration uses: batch size 16, learning rate 10^{-5}, 30 epochs, dropout 0.5.

2.3.3 U-Net segmentation model

The U-Net follows a standard encoder–decoder design with four downsampling and four upsampling stages and skip connections, producing a per-pixel probability map via a sigmoid output activation. Hyperparameter optimisation proceeded in three stages: loss function comparison (Dice vs. cross-entropy), batch size tuning, and joint learning rate–epoch optimisation. Optimal epoch counts were identified from overfitting curves. The final configuration uses: batch size 16, learning rate 10^{-4}, 25 epochs, Dice loss.

2.4 3D CNN

The Spatiotemporal3DCNN is a hybrid 3D + 2D architecture. Three 3D convolutional layers (1→16→32→64 channels) with interleaved max-pooling extract spatiotemporal features, followed by a residual 3D block. The temporal dimension is then collapsed by mean-pooling, and two 2D convolutional stages with a residual-style skip connection refine the spatial representation. Bilinear upsampling restores the output to 288 × 288 pixels, and a final 1 × 1 convolution produces the segmentation map. Due to high computational cost (6 days of training per run on available hardware), no systematic hyperparameter optimisation was performed.

2.5 Proof-of-Concept Clinical Study: In-Vivo Acquisition of Dynamic Infrared Thermography Data for Skin Lesion Characterisation

In the European Union, medical devices are classified under the Medical Device Regulation (MDR, Regulation (EU) 2017/745) into four risk-based categories. Non-invasive active diagnostic devices are classified as Class IIa, moderate risk, diagnostic intent, and provide a relevant regulatory precedent for HypIRskin, which shares a comparable non-invasive, non-ionising mode of action. Conformity assessment for Class IIa devices requires Notified Body involvement and typically incurs costs of €50,000–€500,000. Before clinical measurements can begin, studies must receive independent ethics committee approval, obtain written informed consent, verify patient eligibility, and pseudonymise all collected data. This section describes the proof-of-concept measurement campaign conducted with HypIRskin, covering the measurement protocol, patient questionnaire, preliminary findings, and the practical challenges encountered in a clinical dermatology setting.

2.5.1 Measurement Procedure

The study was conducted in accordance with the Declaration of Helsinki. Ethics committee approval was obtained prior to patient contact. All participants received written study information, had the opportunity to ask questions, and provided written informed consent before enrolment. Eligibility was verified by the attending dermatologist. All collected data were pseudonymised; identifiable information was accessible only to authorised personnel. Measurements were performed in a dermatology department using the standardised HypIRskin three-phase protocol: a steady-state baseline phase, cryogenic cooling with the Zimmer Cryo C device (level 5), and a natural reheating phase during which synchronised RGB and infrared sequences were recorded. Patients were allowed an acclimatisation period before measurement to stabilise skin temperature to room conditions (approximately 21 ℃). Total measurement time was approximately 3 minutes per lesion. After each measurement, patients rated their overall experience on a 0–10 scale and answered items on general discomfort, cooling-related discomfort, and difficulty maintaining a stable position.

3. Results

3.1 HypIRskin

The HypIRskin device was tested in two clinical environments: a dermatology practice and a hospital. Measurements were acquired on confirmed melanoma lesions following the three-phase protocol (steady state → cryogenic cooling → natural reheating). The total measurement time per lesion was approximately three minutes. Data acquired five seconds into the reheating phase showed clear thermal contrast between the lesion and surrounding healthy tissue. The lesion centre temperature recovered more rapidly and reached a higher level than that of the peri-lesional skin, consistent with the elevated metabolic heat generation and blood perfusion modelled in the FEM simulation framework. The maximum observed temperature differential between the malignant lesion and healthy skin was approximately 0.8 ℃, providing a measurable and reproducible signal for lesion characterisation. Figure 5 shows a measurement of a melanoma with the thermal recovery curves. Temperature recovery curves were extracted at two reference points: the lesion centre and the point on the skin surface at the same initial temperature as the lesion centre that is maximally distant from it, serving as the healthy tissue reference, with noise mitigation achieved by averaging over the five nearest neighbouring nodes at each reference location. The shapes of these curves and the persistent temperature differential during reheating are consistent with the thermographic signatures predicted by the FEM tumour study, supporting the validity of the simulation-based analysis framework.

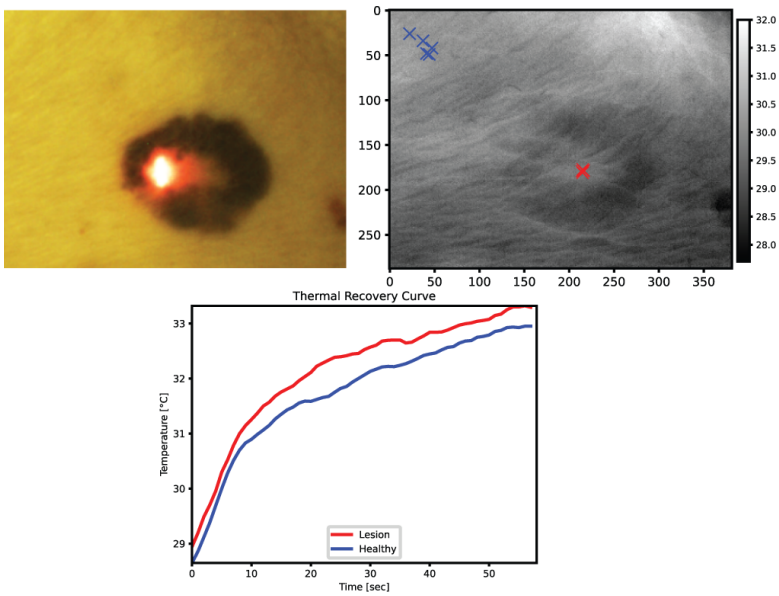


Fig. 5. *Top left: RGB image of the lesion, Top right: corresponding thermogram with five points at the same temperature in center lesion (red) and healthy skin (blue), bottom: Thermal recovery curves of the red (lesion) points and the blue (healthy skin) points.*

3.2 Deep learning

3.2.1 CNN Regression model

The optimised regression CNN achieved a diameter prediction accuracy of 0.200 mm (MSE: 0.0401) and a depth accuracy of 0.117 mm (MSE: 0.0139) on the test set, with a training time of 199 seconds (Table 2). Validation losses converged smoothly over 30 epochs without overfitting.

Table 2. *CNN regression model performance on the test set.*

Training time (s)	Diameter MSE	Depth MSE	Diameter accuracy (mm)	Depth accuracy (mm)
198.83	0.0401	0.0139	0.200	0.117

3.2.2 U-Net segmentation model

The optimised U-Net produced an average absolute area error of 0.434 mm², an average percentage error of 3.68%, and a median percentage error of 0.83% (Table 3). Performance was strong for lesions with diameters $\leq$ 2 mm but degraded markedly for 1 mm diameter tumours and very shallow lesions (depth 0.05 mm), as reflected in the wide mean–median divergence.

Table 3. *U-Net segmentation model performance on the test set.*

Training time (s)	Avg. absolute error (mm^2)	Avg. percentage error (%)	Median percentage error (%)
66,648	0.434	3.68	0.83

3.2.3 3D CNN

Without hyperparameter optimisation, the 3D CNN achieved a test loss below 0.05 after training on 462 lesion geometries, with qualitative segmentation outputs showing good agreement with ground-truth boundaries across a range of diameter and depth combinations.

3.3 Clinical Study

In total, 32 skin lesions from 16 patients were measured. Clinicopathological data from histopathology were available for 20 lesions. The confirmed diagnoses are summarised in Table 4; the cohort was dominated by basal cell carcinomas (BCC), with only three confirmed melanomas.

Table 4. *Clinicopathological composition of the 20 lesions with confirmed diagnoses.*

Diagnosis	N
(Nodular) Basal Cell Carcinoma	10
Melanoma	3
Seborrhoeic keratosis	2
Atypical naevus	1
Keratoacanthoma	1
Morbus Bowen	1
Actinic keratosis	1
Dermal naevus	1
Total	**20**

All three melanoma cases showed a discernible thermal contrast during reheating, with the lesion centre recovering faster and to a higher temperature than the surrounding healthy skin, consistent with the elevated blood perfusion and metabolic heat generation of malignant tissue predicted by the FEM model. Thermographic and RGB images were acquired for each case; no co-registration or preprocessing was applied at this stage.

4. Discussion

4.1 HypIRskin

HypIRskin demonstrates that a compact, integrated, and clinically deployable DIRT system can be constructed around a standardised acquisition protocol, controlled cryogenic excitation, and fully automated software control. The device's reproducibility, short measurement cycle, and non-contact cooling make it a viable candidate for use in dermatological practice, whether as a pre-surgical characterisation tool or as a non-invasive screening adjunct. Several areas for improvement have been identified. The distance holder, while effective, requires design refinement to simplify clinical positioning. A larger patient dataset is needed to establish the optimal cooling settings for different body sites and lesion types, and to enable a statistically grounded evaluation of diagnostic performance. Future software releases will incorporate real-time automated lesion detection and quantitative estimation of subsurface lesion depth and borders using the algorithms developed in the simulation studies. Longer-term development envisions the integration of multispectral or hyperspectral imaging capabilities into the measurement head, providing complementary information on tissue composition and potentially enabling differentiation between malignant and benign lesions without biopsy.

4.2 Deep learning

The regression CNN predicted tumour diameter and depth at 0.200 mm and 0.117 mm accuracy respectively, approximately 25 to 100 times more precise than the coarsest clinical excision margin (5 mm for melanoma in situ). Compared to the FEM-based numerical estimation (errors $\leq$ 0.1 mm for diameter, $\leq$ 0.05 mm for depth on noise-free data), the CNN shows a modest accuracy degradation attributable to the augmentations and noise injected during training to simulate real-world variability, confirming its practical robustness. The U-Net achieved a median area error below 1% across most of the lesion parameter space, but performed poorly for the smallest lesions (1 mm diameter, 0.05 mm depth). Two factors are likely responsible: inaccuracies in the pixel-to-millimetre resolution constant propagate disproportionately for small ground-truth masks, and the U-Net depth was not optimised, a deeper architecture may better resolve the subtle thermal contrast of small or shallow lesions. The 3D CNN, despite lacking any tuning, achieved competitive results, and its native capacity to model the full reheating trajectory without prior dimensionality reduction gives it clear potential to outperform both other models once optimised. A key limitation shared by all three models is that training was conducted exclusively on synthetic FEM data. The idealised layer geometry, isotropic thermophysical properties, and absence of patient-specific variability mean that generalisation to real in-vivo thermograms cannot be assumed. Augmentation partially mitigates this, but prospective clinical validation with histopathologically confirmed Breslow depths is an essential next step. Future work should also retrain on the extended spline-based lesion dataset (Section 2.2) to improve robustness to irregular lesion morphologies, reduce the computational footprint of the 3D CNN, and integrate model outputs directly into the HypIRskin acquisition software for real-time clinical decision support.

4.3 Clinical Study

The proof-of-concept campaign confirms that DIRT measurements can be performed in an active clinical setting with the HypIRskin device, with a per-lesion acquisition time of approximately 3 minutes and no adverse events. The thermal contrast observed in all three confirmed melanoma cases is consistent with FEM predictions, providing preliminary in-vivo support for the physical basis of the approach. The primary limitation is dataset size: 20 lesions with confirmed diagnoses, of which only 3 are melanomas, is insufficient to train or validate the deep learning models described in Section 5. Published precedents for Class IIa approval underscore the scale of data collection still required. Dataset growth was impeded by two factors: protracted ethics approval procedures, compounded by COVID-19 pandemic restrictions on hospital access and patient contact; and the logistical complexity of integrating the measurement protocol into the routine workflow of an active dermatology department, where staff availability and scheduling constraints limited session throughput. Future work will focus on expanding the patient dataset through a dedicated measurement campaign, with priority given to increasing the number of melanoma, atypical naevus, and dermal naevus cases. As sufficient data are accumulated, the synthetic-data-trained models will be fine-tuned on real patient thermograms via transfer learning. Longer-term goals include automated lesion detection, quantitative depth estimation, and a diagnostic distinction between benign and malignant lesions, positioning HypIRskin as a non-invasive complement to dermoscopy and histopathology in clinical decision-making.

Informed Consent

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the University Hospital Antwerp (B300201941125, date of approval: 16 May 2022).

Acknowledgment

The successful design and construction of the active thermography device for skin cancer diagnosis would not have been possible without the invaluable support and collaboration of ZHAW and Mathias Bonmarin.

Funding

This research was funded by the Fonds Wetenschappelijk Onderzoek (FWO) via support for the FWO research project, “Optimized skin tissue identification by combined thermal and hyperspectral imaging methodology.” and the finishing grant "Emmanuel van der Schueren" supported by "Kom Op Tegen Kanker". (project number 41882 (FWO G0A9720N))

4.4 References

References

[1] World Health Organization. *WORLD CANCER REPORT: Cancer Research for Cancer Development.* IARC, Lyon, France, 2020.

[2] Hyuna Sung, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3):209–249, 2021.

[3] Nora Eisemann, Annika Waldmann, Alan C. Geller, Martin A. Weinstock, Beate Volkmer, Ruediger Greinert, Eckhard W. Breitbart, and Alexander Katalinic. Non-Melanoma Skin Cancer Incidence and Impact of Skin Cancer Screening on Incidence. *Journal of Investigative Dermatology*, 134(1):43–50, January 2014.

[4] Mariah M Johnson, Sancy A Leachman, Lisa G Aspinwall, Lee D Cranmer, Clara Curiel-Lewandrowski, Vernon K Sondak, Clara E Stemwedel, Susan M Swetter, John Vetto, Tawnya Bowles, Robert P Dellavalle, Larisa J Geskin, Douglas Grossman, Kenneth F Grossmann, Jason E Hawkes, Joanne M Jeter, Caroline C Kim, John M Kirkwood, Aaron R Mangold, Frank Meyskens, Michael E Ming, Kelly C Nelson, Michael Piepkorn, Brian P Pollack, June K Robinson, Arthur J Sober, Shannon Trotter, Suraj S Venna, Sanjiv Agarwala, Rhoda Alani, Bruce Averbook, Anna Bar, Mirna Becevic, Neil Box, William E Carson, Pamela B Cassidy, Suephy C Chen, Emily Y Chu, Darrel L Ellis, Laura K Ferris, David E Fisher, Kari Kendra, David H Lawson, Philip D Leming, Kim A Margolin, Svetomir Markovic, Mary C Martini, Debbie Miller, Debjani Sahni, William H Sharfman, Jennifer Stein, Alexander J Stratigos, Ahmad Tarhini, Matthew H Taylor, Oliver J Wisco, and Michael K Wong. Skin cancer screening: Recommendations for data-driven screening guidelines and a review of the US Preventive Services Task Force controversy. *Melanoma Management*, 4(1):13–37, March 2017.

[5] Franz Nachbar, Wilhelm Stolz, Tanja Merkle, Armand B. Cognetta, Thomas Vogt, Michael Landthaler, Peter Bilek, Otto Braun-Falco, and Gerd Plewig. The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 30(4):551–559, April 1994.

[6] Alicia Brunssen, Annika Waldmann, Nora Eisemann, and Alexander Katalinic. Impact of skin cancer screening and secondary prevention campaigns on skin cancer incidence and mortality: A systematic review. *Journal of the American Academy of Dermatology*, 76(1):129–139.e10, January 2017.

[7] Jan Verstockt, Simon Verspeek, Filip Thiessen, Wiebren A. Tjalma, Lieve Brochez, and Gunther Steenackers. Skin Cancer Detection Using Infrared Thermography: Measurement Setup, Procedure and Equipment. *Sensors*, 22(9):3327, January 2022.

[8] Jeffrey E. Gershenwald, Richard A. Scolyer, Kenneth R. Hess, Vernon K. Sondak, Georgina V. Long, Merrick I. Ross, Alexander J. Lazar, Mark B. Faries, John M. Kirkwood, Grant A. McArthur, Lauren E. Haydu, Alexander M. M. Eggermont, Keith T. Flaherty, Charles M. Balch, John F. Thompson, and for members of the American Joint Committee on Cancer Melanoma Expert Panel and the International Melanoma Database and Discovery Platform. Melanoma staging: Evidence-based changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA: A Cancer Journal for Clinicians*, 67(6):472–492, 2017.

[9] Donald L. Morton, David G. Davtyan, Leslie A. Wanek, Leland J. Foshag, and Alistar J. Cochran. Multivariate analysis of the relationship between survival and the microstage of primary melanoma by clark level and breslow thickness. *Cancer*, 71(11):3737–3743, 1993.

[10] Ayca Kirimtat, Ondrej Krejcar, and Ali Selamat. A Mini-review of Biomedical Infrared Thermography (B-IRT). In Ignacio Rojas, Olga Valenzuela, Fernando Rojas, and Francisco Ortuño, editors, *Bioinformatics and Biomedical Engineering*, volume 11466, pages 99–110. Springer International Publishing, Cham, 2019.

[11] J. H. Flores-Sahagun, J. V. C. Vargas, and F. A. Mulinari-Brenner. Analysis and diagnosis of basal cell carcinoma (BCC) via infrared imaging. *Infrared Physics & Technology*, 54(5):367–378, September 2011.

[12] Haolong Chen, Kaijie Wang, Zhibo Du, Weiming Liu, and Zhanli Liu. Predicting the thermophysical properties of skin tumor based on the surface temperature and deep learning. *International Journal of Heat and Mass Transfer*, 180:121804, December 2021.

[13] M. Pirtini Çetingül, Rhoda M. Alani, and Cila Herman. Quantitative Evaluation of Skin Lesions Using Transient Thermal Imaging. In *2010 14th International Heat Transfer Conference, Volume 1*, pages 31–39, Washington, DC, USA, January 2010. ASMEDC.

[14] Muge Pirtini Cetingul and C. Herman. The Assessment of Melanoma Risk Using the Dynamic Infrared Imaging Technique. *Journal of Thermal Science and Engineering Applications*, 3(3), September 2011.

[15] J. Iljaž, L. C. Wrobel, M. Hriberšek, and J. Marn. Numerical modelling of skin tumour tissue with temperature-dependent properties for dynamic thermography. *Computers in Biology and Medicine*, 112, September 2019.

[16] Müge Pirtini Çetingül, Rhoda M. Alani, and Cila Herman. Detection of Skin Cancer Using Transient/Thermal Imaging. In *ASME 2010 Summer Bioengineering Conference, Parts A and B*, pages 601–602, Naples, Florida, USA, June 2010. American Society of Mechanical Engineers.

[17] Chang W. Song, Anna Lokshina, Juong G. Rhee, Marsha Patten, and Seymour H. Levitt. Implication of Blood Flow in Hyperthermic Treatment of Tumors. *IEEE Transactions on Biomedical Engineering*, BME-31(1):9–16, January 1984.

[18] M. Agrawal, K. R. Pardasani, and Neeru Adlakha. Finite element model to study the thermal effect of tumors in dermal regions of irregular tapered shaped human limbs. *International Journal of Thermal Sciences*, 98:287–295, December 2015.

[19] H. H. Pennes. Analysis of tissue and arterial blood temperatures in the resting human forearm. 1948. *Journal of applied physiology*, 1948.

[20] M. Pirtini Çetingül and C. Herman. Quantification of the thermal signature of a melanoma lesion. *International Journal of Thermal Sciences*, 50(4):421–431, April 2011.

[21] Jan Verstockt, Warre Clarys, Michaël Hillen, Simon Verspeek, Mathias Bonmarin, Filip E. F. Thiessen, Lieve Brochez, and Steenackers Gunther. HypIRskin: Thermography-Guided Device for Diagnosis and Characterization of Skin Cancer Lesions. In *2024 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–6, June 2024.