POLITECNICO DI TORINO
Repository ISTITUZIONALE

U-Net convolutional architecture for squealer tip film cooling prediction and multi-objective optimization

Original
U-Net convolutional architecture for squealer tip film cooling prediction and multi-objective optimization / Orlando, D.A., Salvadori, S., Misul, D.A.. - In: ENERGY AND AI. - ISSN 2666-5468. - ELETTRONICO. - 26:(2026), pp. 1-20. [10.1016/j.egyai.2026.100899]

Availability:
This version is available at: 11583/3015411 since: 2026-09-13T10:55:27Z

Publisher:
Elsevier

Published
DOI:10.1016/j.egyai.2026.100899

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

26 November 2026

Energy and AI 26 (2026) 100899

Contents lists available at ScienceDirect

Energy and AI

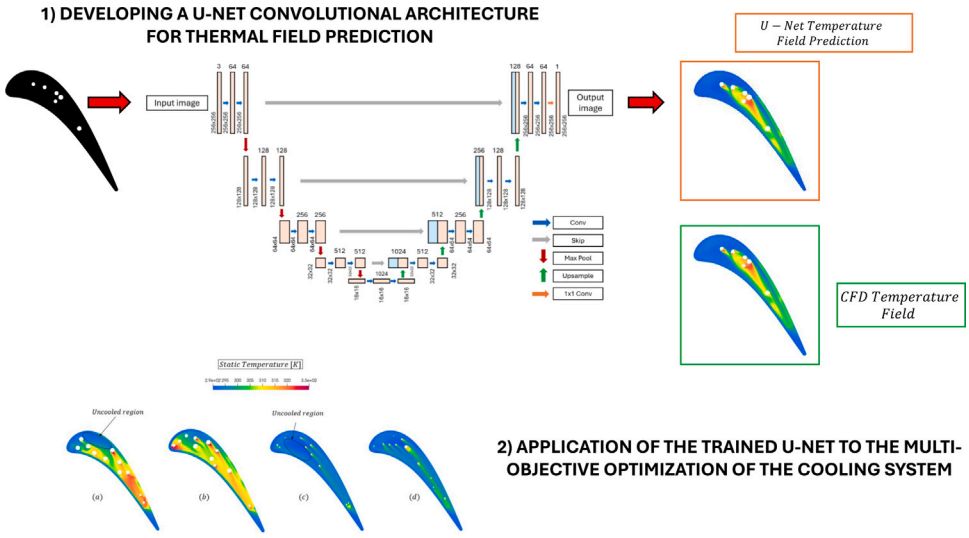
journal homepage: www.elsevier.com/locate/egyai

U-Net convolutional architecture for squealer tip film cooling prediction and multi-objective optimization

Dario Antonio Orlando, Simone Salvadori *, Daniela Anna Misul

Polytechnic University of Turin, Department of Energy, Turin, 10129, Italy

GRAPHICAL ABSTRACT



HIGHLIGHTS

- A CFD–ANN–U-Net framework optimizes blade-tip thermal load.
- U-Net reconstructs temperature fields from cooling-hole layouts.
- U-Net exploits spatial temperature maps to achieve effective cooling-hole layouts.
- This framework reduces computational costs with respect to CFD-based optimizations.

ARTICLE INFO

Keywords:
Artificial Neural Network
Convolutional Neural Network
Optimization
Squealer tip
Film cooling

ABSTRACT

The optimization of turbine blade squealer tip cooling systems is limited by the high computational cost of Computational Fluid Dynamics simulations, which restricts the exploration of large design spaces and the identification of efficient cooling layouts. This work proposes a machine-learning-based framework for the cooling system of the $GE - E^3$ high-pressure turbine blade tip, with fixed squealer cavity geometry and variable number, diameter, and position of film cooling holes. A dataset of 2000 Computational Fluid Dynamics simulations is used to train two Artificial Neural Networks for predicting area-averaged adiabatic effectiveness

* Corresponding author.
E-mail address: simone.salvadori@polito.it (S. Salvadori).

https://doi.org/10.1016/j.egyai.2026.100899
Received 2 July 2026; Received in revised form 28 August 2026; Accepted 5 September 2026
Available online 12 September 2026
2666-5468/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

and coolant mass flow rate, and a U-Net for reconstructing the full tip temperature field. The trained surrogate models are coupled with a Particle Swarm Optimization algorithm using four objective functions representing different cooling design priorities. The framework rapidly identifies optimized configurations that vary according to the selected objective. The best solution simultaneously improves average tip adiabatic effectiveness, coolant mass flow rate, and temperature-field deviation from the coolant temperature relative to the baseline design. The U-Net reconstructs the full temperature field with maximum prediction errors below approximately 5.92 kelvin, while reducing design-evaluation time from several hours to a few seconds. For the investigated geometry and operating conditions, the optimized layouts indicate that coolant is more effectively distributed through multiple small-diameter holes, whereas holes located in high-pressure tip regions may provide limited benefit because of the local surface pressure. These results show that data-driven surrogate models combined with evolutionary optimization can enable rapid and effective exploration of turbine blade cooling designs.

1. Introduction

In turbine blades, the design of the squealer tip is an extremely important aspect to take into account, both for aerodynamic and thermal considerations. In fact, the tip gap accounts for about one-third of the aerodynamic losses that occur in the rotor [1]. Furthermore, the hot flows originating from the combustor tend to subject the tip to extremely high stress due to the vortical structures that scrape and impinge upon that surface, drastically impacting the life span of the component [2,3].

The thermal management of modern gas turbines relies on the combined development of external film cooling and internal cooling systems. A comprehensive review by Zhang et al. [4] summarized the main optimization strategies adopted for gas turbine cooling structures, including both film cooling configurations and internal cooling passages. For the latter, geometric features such as ribs and dimples are widely employed to enhance heat transfer, although their performance is strongly coupled with rotation and the associated flow structures. Che et al. [5], for example, investigated the influence of dimple density and diameter on heat transfer and flow resistance in rotating cooling channels, while You et al. [6] analyzed the Coriolis-induced asymmetry of the drag mechanisms in rotating ribbed channels. These studies highlight the strong interaction between cooling system geometry, rotational effects, heat transfer, and aerodynamic losses, motivating the development of systematic optimization strategies for turbine cooling systems.

To successfully optimize the tip geometry, numerous approaches have been proposed over the years, which include both the cooling holes and the squealer geometry in the optimization process. Andreoli et al. [7] performed an evolutionary multi-objective optimization of the tip geometry of a cooled turbine blade in order to optimize both the heat load and the aerodynamic efficiency. Specifically, the number of holes, their diameter, inclination, and position were kept constant, but a 44-parameter geometric parametrization allowed for the arbitrary modification of the rim shapes around the coolant holes, which are arranged both on the tip surface and on the pressure side. To reduce the computational cost, only the rotor was solved, inserting the boundary conditions obtained from a previous simulation that included the vane at the inlet, and Reynolds-Averaged Navier–Stokes equations (RANS) were used. The optimized geometry allowed for a 65% reduction in heat flux and a 0.53% improvement in efficiency compared to the baseline configuration at a constant coolant flow rate, thus improving the mixing processes due to better rim-coolant interaction. Specifically, it is shown how the tip holes modify the over-tip flow field, thus impacting both the thermal and aerodynamic performance. De Maesschalck et al. [8] proposed a more comprehensive optimization strategy that not only optimizes the rim geometry of the squealer tip, but also includes the positioning, the number, and the diameter of the cooling holes. A specific parametrization was used to ensure no intersection between the rim geometry and the cooling holes. This method significantly increases the extent of the design space with 40 design variables. In this case, both aerodynamic efficiency and heat load aspects were taken into account using a RANS model. The parametrization is highly innovative and involves a cubic interpolated curves approach for both the hole positions and the rim geometry. The mechanism includes the hole or the rim in areas where the interpolation curve exceeds a certain threshold and excludes them below that threshold. This is achieved through a rectangular control surface which is then re-mapped onto the blade profile for the rim and through 3 pre-defined lines for the cooling holes. Optimizations were performed for both a cooled and an uncooled case. The optimization for uncooled squealer tips yielded simpler, highly efficient double-rim profiles, favoring either an open leading edge combined with an open suction side rim or an unclosed rear pressure side. For cooled geometries, the Pareto front showed designs that significantly reduced metal temperatures while maintaining aerodynamic efficiency, featuring common elements like enhanced leading edge cooling and a hook shape. Duan et al. [9] employed a genetic algorithm to optimize a multi-objective function that included temperature, temperature variance, and aerodynamic efficiency by varying both the cooling hole configuration (where 32 fixed slots are parameterized such that in every configuration 10 holes are active and 22 are inactive) and the squealer tip height. They compared a conventional sequential optimization for squealer tip shaping and cooling holes optimization with a concurrent combined optimization of both aspects. The results demonstrated that, unlike classic film cooling, the tip cooling system exhibits a strong interaction with the external flow, particularly the Tip Leakage Vortex (TLV). The ejected coolant creates a blockage effect that effectively reduces aerodynamic losses. This study confirmed that the concurrent optimization approach is superior to the conventional sequential method, as the latter can fail to capture significant potential performance gains. In general, increasing the squealer height was found to improve both thermal and aerodynamic characteristics by enhancing the blockage effect and trapping the coolant flow within the tip region. Guo et al. [10] optimized a tip cooling system using oval-shaped holes, parameterizing their position (within the vicinity of a reference hole location) and their inclination. RANS modeling was adopted, and a global optimization algorithm based on a Kriging surrogate model was used. They showed that a positive axial inclination of the holes and a closer spacing of the holes improved the distribution of the film cooling effectiveness. The positive axial inclination also enhanced the blockage effect on the leakage flow. While classic cylindrical holes achieved a 54.85% improvement in adiabatic film cooling effectiveness, the use of oval holes reached a 67.65% improvement. Caloni et al. [11], on the other hand, performed a Multi-Disciplinary Design Optimization (MDO), carrying out both a Conjugate Heat Transfer (CHT) analysis and a Finite Element (FE) analysis to evaluate the stress the blade must endure. They optimized both the aerodynamic performance and the stress level of the blade. The squealer geometry was parameterized to allow for variations in its height and thickness at different points. Furthermore, two of the three holes at the tip could move along the chamber line, and the holes exiting at the pressure side could vary their inclination. The optimization problem was solved using a Multipoint Approximation Method, which resulted in an improvement of 0.13% in aerodynamic efficiency and 5% in maximum stress by modifying the squealer geometry.

Nomenclature

Latin symbols	
d_i	Cooling-hole diameter, mm
$\dot{m}_{\text{cool}}$	Coolant mass flow rate, kg s^{-1}
N_Ω	Number of pixels in the physical tip domain
$p_{0,\text{cool}}$	Coolant total pressure, Pa
$p_{0,\text{main}}$	Mainstream total pressure, Pa
q''	Wall-normal heat flux, W m^{-2}
R^2	Coefficient of determination
$\text{RMSD}(T, T_{0,\text{cool}})$	Root-mean-square temperature deviation, K
T_i	Temperature at pixel i, K
$T_{0,\text{cool}}$	Coolant total temperature, K
$T_{0,\text{main}}$	Mainstream total temperature, K
T_{aw}	Adiabatic-wall temperature, K
$\overline{\overline{T_{\text{tip}}}}$	Area-averaged tip temperature, K
x_i	Streamwise coordinate of cooling hole i, mm
y_i	Transverse coordinate of cooling hole i, mm
$y_{\text{SS}}, y_{\text{PS}}$	Limits of the lateral averaging interval
y^+	Dimensionless wall distance
Greek symbols	
α_{in}	Mainstream inlet flow angle, °
η_{aw}	Local adiabatic film cooling effectiveness
$\overline{\eta_{\text{aw}}}$	Laterally averaged adiabatic effectiveness
$\overline{\overline{\eta_{\text{aw}}}}$	Area-averaged tip adiabatic effectiveness
Ω	Set of pixels belonging to the physical tip domain
Acronyms	
ANN	Artificial Neural Network
BR	Blowing Ratio
CAD	Computer-Aided Design
CFD	Computational Fluid Dynamics
CHT	Conjugate Heat Transfer
CNN	Convolutional Neural Network
DoE	Design of Experiments
HPC	High-Performance Computing
LHS	Latin Hypercube Sampling
LUT	Look-Up Table
MAE	Mean Absolute Error
MSE	Mean Squared Error
NCM	Nearest Color Mapping
PSO	Particle Swarm Optimization
RANS	Reynolds-Averaged Navier–Stokes
RGB	Red, Green, Blue
RMSD	Root-Mean-Square Deviation
Subscripts	
aw	Adiabatic wall
cool	Coolant
in	Inlet
main	Mainstream
norm	Normalized quantity
PS	Pressure side
SS	Suction side
tip	Blade tip

However, these optimization approaches have the disadvantage of severely limiting the design exploration space due to the limited number of simulations that can be performed, given the computational cost of computational fluid dynamics (CFD) models. Similar challenges associated with the analysis and optimization of strongly coupled, multi-parameter systems are encountered throughout the broader energy and materials fields, including energy-storage ceramics, photocatalytic materials, electrochemical energy-storage systems, and renewable-energy management [12–15]. This broader context further motivates the development of computationally efficient predictive and optimization methodologies. To address this problem in fluid and thermal engineering, a growing number of studies adopting machine learning (ML) methods are emerging, often allowing for a better capture of the non-linearity between input and output compared to more classical surrogate models.

Beyond conventional purely data-driven models, physics-informed neural networks (PINNs) have also demonstrated the possibility of incorporating governing physical laws directly into the learning process by embedding differential-equation residuals within the training procedure [16]. Furthermore, more advanced deep-learning methods allow for considering as output an entire temperature or film cooling effectiveness map instead of an averaged value, thus improving the spatial quality of the analysis.

Recent studies have specifically demonstrated this potential for film cooling applications. Huang et al. [17] developed a Conditional Generative Adversarial Network (CGAN) to predict the two-dimensional distribution of adiabatic film cooling effectiveness for a mist/air cooling configuration, demonstrating the capability of image-based deep-learning models to reproduce complex spatial cooling patterns at a computational cost substantially lower than CFD. Similarly, Liu et al. [18] investigated the parameter-coupling mechanisms of a serrated-V-shaped film cooling configuration and employed a Generalized Regression Neural Network (GRNN) to predict its cooling effectiveness, further demonstrating the applicability of ML surrogate models to highly nonlinear film cooling problems.

Tao et al. [19] used an improved image-to-image translation Generative Adversarial Network (GAN) as a surrogate model to improve both the aerodynamic efficiency and the adiabatic film cooling effectiveness, thus using the Non-dominated Sorting Genetic Algorithm III (NSGA-III) as the optimizer. This approach allowed for achieving a high-prediction reliability, with a Mean Absolute Error (MAE) of less than 0.15% and a correlation index $R^2 = 0.999$. In this optimization, the characteristics of the cooling system were maintained unchanged, while the cavity bottom blade tip was modified, showing a noticeable improvement in film cooling attachment and thus improving coverage. It is also observed that for this geometry, an increase in cavity height does result in an increase in film cooling effectiveness, but it penalizes aerodynamic efficiency. Labrini et al. applied an image-based CNN directly within the aerodynamic optimization of a high-pressure turbine vane designed for pressure-gain combustion cycles [20]. The CNN estimated vane performance from images encoding the blade and diffusive-endwall geometries and was coupled with a Particle Swarm Optimization algorithm to explore an 18-variable design space generated through RANS simulations. The resulting optimized profiles improved the vane efficiency relative to the nominal configuration and retained robust performance under pulsating inlet conditions representative of a rotating detonation combustor, while also attenuating the imposed pressure oscillations. This work demonstrates how CNN surrogates can replace repeated CFD evaluations within turbomachinery optimization while preserving the geometric information contained in image-based inputs. Moreover, Yildiz et al. [21] utilized an Artificial Neural Network (ANN) to optimize the layout of film cooling holes for two different blowing ratios (BRs), BR = 0.4 and BR = 1.0. A multi-objective optimization resulted in a 13.43% reduction in total heat flux and a 0.4% improvement in aerodynamic efficiency compared to the baseline case.

This article develops a data-driven framework that combines scalar and full-field neural-network models with an optimization algorithm for the design of a squealer-tip cooling system. Although U-Net/GAN-based models have already been coupled with optimization algorithms in film cooling studies, the main contributions of the present work are summarized below:

- To the best of the authors' knowledge, this is the first use of a U-Net within an optimization loop to directly redesign the cooling

system of a squealer tip. Unlike the GAN-based approach previously described in [19], the present framework directly optimizes the activation, diameter, and position of the cooling holes while keeping the squealer-tip geometry fixed. The image-based model is therefore used to explore and redesign the cooling layout itself.

- A geometry-dependent mask is introduced in both U-Net training and design evaluation. Although masking is not new in image-based learning, to the best of the authors' knowledge, it has not previously been applied to film cooling field prediction. This is important because the squealer-tip surface is irregular and its physical domain changes with the cooling-hole layout. The mask excludes the image background and hole interiors from the loss and performance metrics, preventing these large and nearly uniform regions from hiding errors over the actual tip surface. It therefore allows different cooling layouts to be treated on the same rectangular image grid while focusing the learning process only on the region of physical interest.
- The ANN and U-Net are compared from an engineering point of view rather than only in terms of prediction accuracy. The aim is to determine whether the higher cost and complexity of the U-Net are justified by the full-field information that it provides. In particular, the analysis assesses whether this additional spatial information leads to different or better optimized layouts and provides useful insights into local coolant coverage and jet interactions, or whether the scalar predictions of the simpler ANN are sufficient for the design problem. The comparison thus provides practical guidance on the level of model complexity that is actually needed. This is relevant in industrial applications, where a more complex method is useful only when it leads to better design decisions or a clearer understanding of the cooling process.
- The 44-variable parametrization provides independent control of the activation, diameter, and two-dimensional position of each cooling hole. The holes can move within the admissible squealer cavity and can be added or removed without being restricted to fixed candidate slots, prescribed lines, or spline-based layouts. This broad design space is combined with criteria based on $\overline{\overline{\eta_{aw}}}$, $\dot{m}_{cool}$, and RMSD$(T, T_{0,\mathrm{cool}})$. To the best of the authors' knowledge, no previous squealer-tip cooling study has combined this level of independent hole control with U-Net field prediction and optimization criteria accounting for cooling performance, coolant consumption, and the spatial temperature field.

2. Test case

The geometry chosen for this optimization study is that of the $GE-E^3$ blade first high-pressure stage. The geometry is obtained through a simple linear extrusion of the tip profile, without incorporating radial twist along the blade span, in order to replicate the experimental conditions of Kwak et al. [22], who conducted the test on a five-bladed linear cascade with a blow-down facility shown in [23].

To simplify the experimental setup, the main flow is maintained at ambient temperature, while the cooling air is heated using an heater. The blade geometry also features 11 film cooling holes arranged along the camber line of the profile, whose axial chord is 8.61 cm, with a distance of 0.635 cm from one to the other and a diameter of 0.127 cm. The distribution of adiabatic film cooling effectiveness is therefore evaluated using a hue detection-based transient liquid crystal technique. Adiabatic data are obtained by coating the blade tip with low-conductivity black polycarbonate. From the distribution on the tip, it is also possible to calculate the distribution along the blade's axial coordinate of the laterally averaged adiabatic film cooling effectiveness as follows:

$$\eta_{\mathrm{aw}} = \frac{T_{\mathrm{m}} - T_{\mathrm{aw}}}{T_{\mathrm{m}} - T_{\mathrm{c}}}, \qquad \overline{\eta_{\mathrm{aw}}}(x) = \frac{1}{y_{\mathrm{PS}} - y_{\mathrm{SS}}} \int_{y_{\mathrm{SS}}}^{y_{\mathrm{PS}}} \eta_{\mathrm{aw}}(x, y)\, \mathrm{d}y. \tag{1}$$

Here, T_{m}, T_{c}, and T_{aw} are the mainstream, coolant, and adiabatic-wall temperatures, respectively. At each axial location, y_{SS} and y_{PS} identify the suction-side and pressure-side limits of the averaging interval. In the computer-aided design (CAD) model, shown in Fig. 1, only one of the 5 blades was included to reduce computational cost, exploiting the periodicity condition of the walls. Furthermore, a plenum was included inside the blade to supply the film cooling holes.

2.1. Computational mesh

The computational domain was discretized using the polyhedral meshing capability available within the ANSYS Fluent Meshing software environment. A global cell size constraint was imposed, limiting the maximum element size to 0.008 m. To ensure optimal near-wall resolution, 20 inflation layers were applied along all wall boundaries, starting with a first cell height of 5×10^{-7} m. This fine resolution, coupled with a constant growth rate of 1.3, was specifically designed to meet the strict requirement of achieving approximately $y^+ \approx 1$ across the entire blade surface, aligning with the needs of the chosen turbulence model. A body of influence (BOI) was strategically placed around the squealer tip region to enhance local mesh quality, targeting an element size of 0.0006 m, crucial for accurately resolving the complex leakage flow vortex and the mixing phenomena. Furthermore, to ensure precise definition of the coolant entry, a dedicated surface mesh control was implemented on the cooling hole walls, enforcing a cell size of 6×10^{-5} m. To confirm that the numerical results were independent of the spatial discretization, an initial mesh independence study was performed. Three distinct mesh resolutions were investigated: a fine grid comprising 6 million elements, a medium grid with 3 million elements, and a coarse grid containing 1.7 million elements.

2.2. Simulation setup and validation

The Computational Fluid Dynamics (CFD) simulations were executed using the commercial software ANSYS Fluent. A pressure-based solver, utilizing a coupled scheme for enhanced stability and convergence, was employed. Spatial discretization of the governing equations was achieved using a robust 2nd order scheme. The solution procedure was terminated after 1500 iterations, ensuring that all defined convergence criteria were fully satisfied. The Realizable k-ε model coupled with the Enhanced Wall Treatment was selected to model the flow turbulence. This specific combination was chosen because, unlike the standard wall function approach used by the classic k-ε model, the Enhanced Wall Treatment resolves the viscous sub-layer directly. Consequently, this required a highly refined mesh near the walls, specifically demanding a non-dimensional distance of the first cell $y^+ \approx 1$ across the entire solid domain.

Specific flow parameters were imposed at the inlet and outlet boundaries:

- Main Flow Inlet: The main flow was defined by a total pressure $p_{0,\mathrm{main}} = 126900$ Pa and a total temperature $T_{0,\mathrm{main}} = 297$ K. The flow was directed at an angle $\alpha_{\mathrm{in}} = 32°$. Turbulence characteristics were specified as a turbulent intensity of 9.7% and a turbulent length scale of 0.015 m.
- Coolant Injection Inlet: For the film cooling flow, a total temperature $T_{0,\mathrm{cool}} = 328$ K and a total pressure $p_{0,\mathrm{cool}} = 113000$ Pa were prescribed. These settings were chosen to achieve the target blowing ratio (BR = 1). The coolant turbulence was characterized by an intensity of 5% and a turbulent viscosity ratio of 10.
- Wall Condition: The solid walls were treated as adiabatic surfaces, corresponding to a zero wall-normal heat-flux condition ($q'' = 0\,\mathrm{W\,m^{-2}}$).

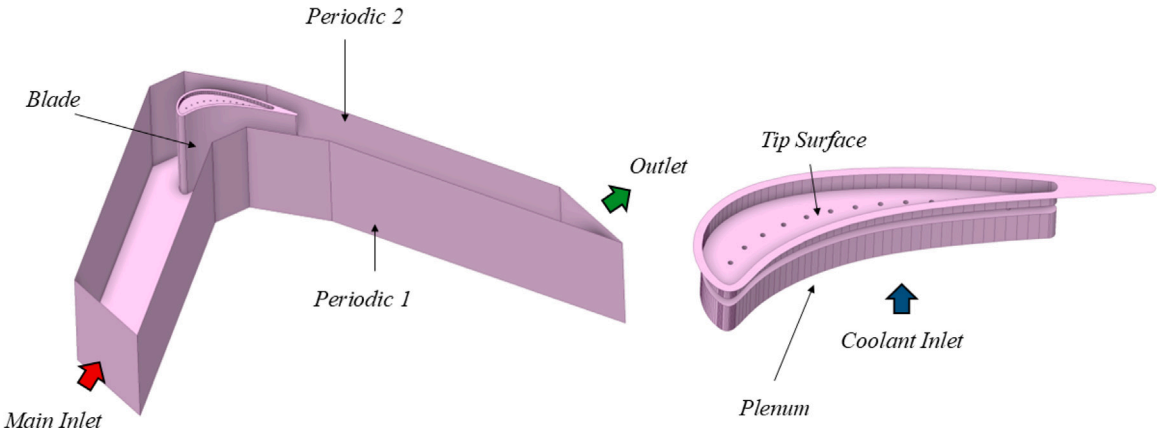


Fig. 1. Computational domain adopted for the CFD simulations, showing the mainstream passage, squealer-tip region, cooling holes, and coolant plenum.

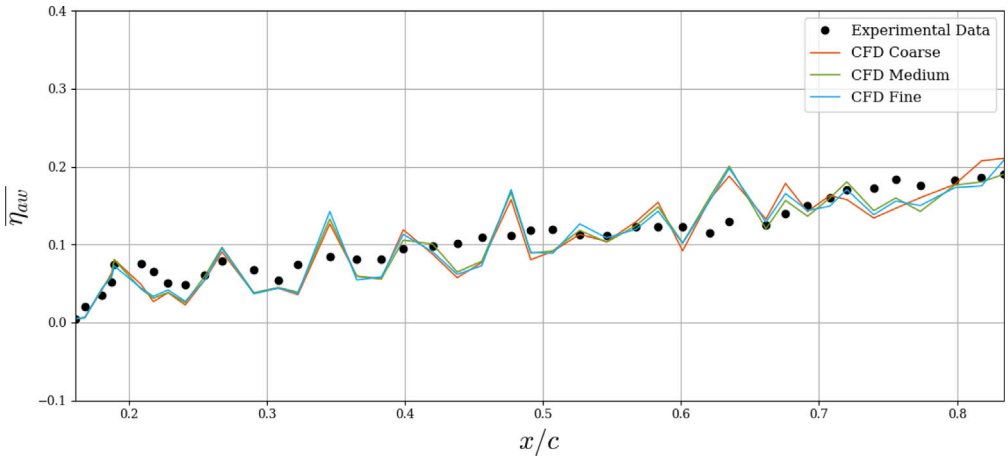


Fig. 2. Comparison of the laterally-averaged adiabatic film cooling effectiveness obtained with the coarse, medium, and fine computational meshes against the available experimental measurements.

The comparison between the obtained results in terms of laterally-averaged adiabatic film cooling effectiveness shown in Fig. 2 demonstrates that the selected numerical setup accurately reproduces the available experimental measurements, thereby validating the present CFD methodology. In particular, the numerical results correctly capture both the overall trend of the effectiveness distribution and the main local peaks along the streamwise direction, with only minor discrepancies in a few localized regions. A detailed discussion of the corresponding baseline flow field and the associated thermal features is provided in Section 4.2. This validation process also confirmed that the selected grid density was sufficient to ensure grid-independent results. Accordingly, the Grid Convergence Index (GCI), computed on the mean tip temperature, was found to be 1.32%. Therefore, the medium mesh (containing 3 million elements, as detailed in the previous section) was adopted for the final uncertainty analysis and all subsequent computations. The execution time for a single CFD simulation, run on 48 CPU cores, is approximately 35 min.

3. Methodology

The proposed workflow replaces repeated CFD evaluations during optimization with surrogate models trained on an initial CFD database. The cooling-hole geometry is first parameterized and sampled through a Sobol sequence. CFD simulations are then performed for the sampled designs, and the resulting scalar and spatial outputs are used to train two types of surrogate model: fully connected Artificial Neural Networks (here referred to as ANNs) for global scalar quantities (like $\dot{m}_{cool}$ and $\overline{\overline{T_{tip}}}$ from which it can then be obtained by simple rescaling also $\overline{\overline{\eta_{aw}}}$) and a U-Net Convolutional Neural Network (CNN) for the complete temperature field. The trained surrogates are subsequently coupled with a Particle Swarm Optimization (PSO) algorithm.

Compared with a conventional parametric optimization, which may require thousands of simulations and may converge toward a local minimum, the trained neural networks can evaluate hundreds of thousands of candidate objective-function values within a few seconds. This enables a much broader exploration of the design space, including solutions that may not be apparent a priori. The same surrogate models can also be used to decompose the influence of individual parameters, perform Sobol sensitivity analyses, and rapidly repeat the optimization when constraints or objective weights are changed.

3.1. Geometric parametrization and design-of-experiments generation

The CAD model was parameterized using 44 design variables, corresponding to four variables for each of the 11 possible cooling holes. For hole i, these variables are the diameter d_i, the horizontal coordinate x_i, the vertical coordinate y_i, and the activation parameter $(\text{On/Off})_i$, which determines whether the hole is present. The hole-center coordinates are defined in the x, y reference system shown in Fig. 3, whose origin is located at the airfoil trailing edge.

The Design of Experiments (DoE) used to generate the training database was obtained through Sobol sampling. Hole diameters were

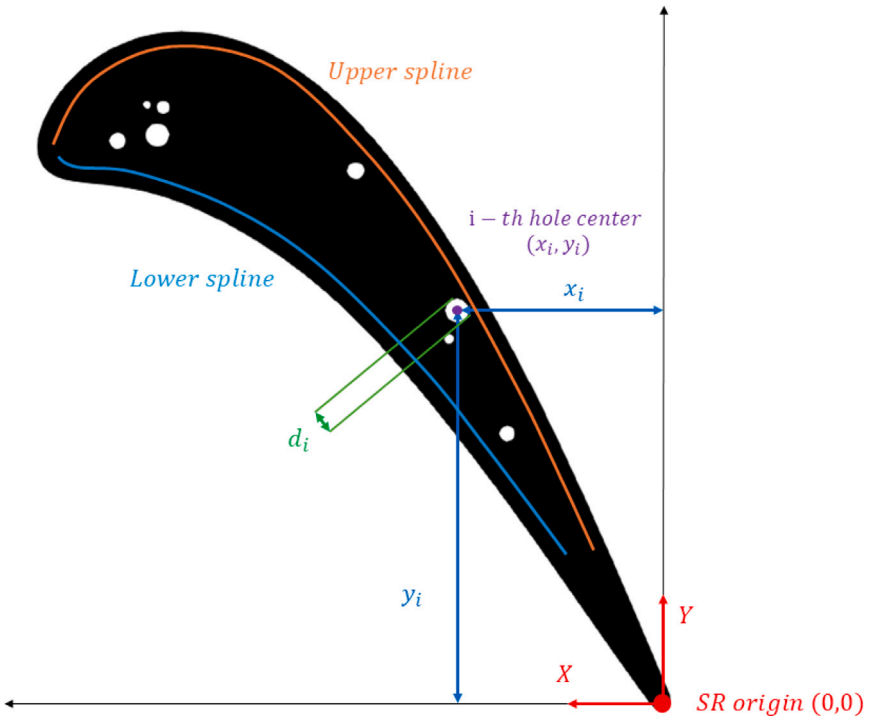


Fig. 3. Geometry-parametrization approach used to define hole position, diameter, and activation within the fixed squealer geometry.

Table 1
Summary of the adopted parametrization.

Characteristic	Value
Maximum Number of Cooling Holes	11
Total Number of Optimization Parameters	44
Number Parameters for Each Cooling Hole	x_i, y_i, d_i, $(\mathrm{On/Off})_i$

Table 2
Bounds of the optimization parameters.

Optimization parameter	Bounds
d_i [mm]	[1 – 4]
x_i [mm]	[22.20 – 86.65]
y_i [mm]	[12.37 – 80.95]
$(\mathrm{On/Off})_i$	[0 *or* 1]

sampled between 1 mm and 4 mm. The activation variables were initially sampled continuously between 0 and 1 and subsequently discretized: values below 0.5 were assigned to 0, whereas values greater than or equal to 0.5 were assigned to 1.

A conditional strategy was adopted for the hole-center coordinates to prevent holes from being placed outside the internal squealer cavity. The horizontal coordinate x_i was sampled freely between 21 mm and 81 mm, corresponding to the left and right cavity limits in the adopted reference system. The vertical coordinate y_i was then sampled conditionally. Two splines, denoted as the upper and lower splines, were constructed along the internal cavity boundaries. Once x_i was selected, the two splines provided the local upper and lower admissible limits for y_i, thereby ensuring that the hole center remained inside the cavity. This strategy preserves a wide design space while maintaining geometrical feasibility. The tip gap and squealer height were kept fixed throughout the parametrization.

To summarize, the main characteristics of the adopted parametrization are presented in Table 1, while the bounds adopted for the 44 optimization parameters are reported in Table 2.

Industrial manufacturability is ensured by restricting the maximum number of cooling holes to that of the baseline configuration and by limiting the hole diameters to a range which is representative of dimensions that can be reliably manufactured using conventional hole-manufacturing techniques. Furthermore, the optimization algorithm presented in Section 3.3 guarantees the feasibility of the generated designs through a penalty function. Specifically, candidate solutions are not penalized only if no overlap exists between active holes and all active holes are entirely contained within the cavity boundaries. In addition, a prescribed positive hole-to-hole and hole-to-wall clearance is included in these checks to preserve the structural integrity of the component.

3.2. Surrogate models

The CFD database provides two types of target quantity. The area-averaged adiabatic film cooling effectiveness $\overline{\overline{\eta_{\mathrm{aw}}}}$ and coolant mass flow rate $\dot{m}_{\mathrm{cool}}$ are global scalar outputs and are predicted using fully connected ANNs. The spatial temperature distribution (from which the distribution of η_{aw} can therefore also be obtained) is instead represented by a 256 × 256 scalar map and is predicted using a U-Net. The following subsections describe the two model families and the preprocessing required for the spatial target.

3.2.1. Artificial neural networks for scalar outputs

Two separate single-output ANNs were used, one for $\overline{\overline{\eta_{\mathrm{aw}}}}$ and one for $\dot{m}_{\mathrm{cool}}$. In both cases, the input vector $\mathbf{x}_{\mathrm{in}}$ contains the same 44 geometric parameters, while the network returns one scalar prediction $\hat{y}_{\mathrm{out}}$. Each architecture contains between one and three fully connected hidden layers.

The networks were trained using the Adam optimizer for a maximum of 2000 epochs, with Mean Squared Error (MSE) as the loss function. All inputs were normalized before training. The dataset was divided into training, validation, and test subsets using a 70%/15%/15% split, and early stopping with a patience of 50 epochs was employed to limit overfitting.

The hyperparameters were selected through a random search over 500 candidate configurations, each evaluated using 5-fold cross-validation. The number of hidden layers was allowed to vary from one

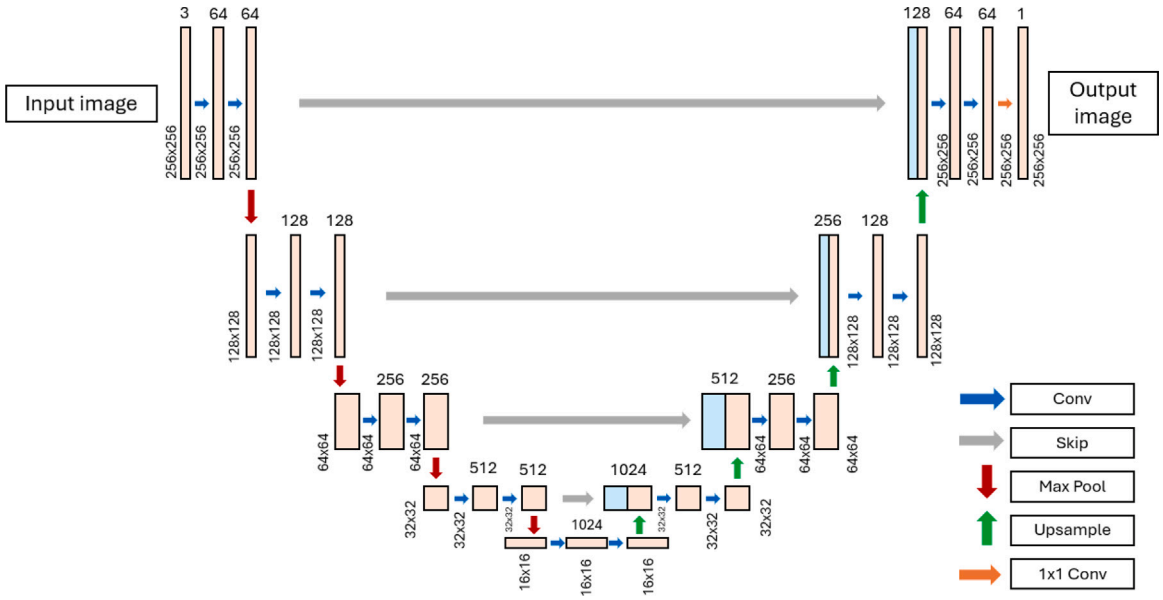


Fig. 4. U-Net encoder–decoder architecture adopted for temperature-field prediction, including skip connections between corresponding resolution levels.

Table 3
Hyperparameter search ranges for the ANN models.

Hyperparameter	Search range
Number of Hidden Layers	[1 – 3]
Hidden Layer 1 Neurons	[8 – 255]
Hidden Layer 2 Neurons	[8 – 255]
Hidden Layer 3 Neurons	[8 – 255]
α_{LR}	$[1 \times 10^{-5} - 1 \times 10^{-3}]$
Batch Size	[10 – 80]
λ_{wd} (weight decay)	$[1 \times 10^{-7} - 1 \times 10^{-4}]$
Dropout Value	[0.05, 0.10, 0.15, 0.20, 0.25]
Activation Function	[ReLU, Tanh, Sigmoid, LeakyReLU]

to three. The remaining search variables were the number of neurons in each layer, activation function, learning rate α_{LR}, batch size, dropout probability, and weight-decay coefficient λ_{wd}. The complete search space is reported in Table 3.

3.2.2. U-Net architecture

Unlike fully connected ANNs, Convolutional Neural Networks (CNNs) are designed to process spatially structured data such as images by extracting local features through trainable convolutional filters [24]. This makes them particularly suitable for problems in which the spatial distribution of the input contains relevant physical information. When the desired output is also a spatial field rather than a single scalar quantity, an encoder–decoder architecture can be used to progressively extract features and subsequently reconstruct the output at the original spatial resolution.

For the present image-to-image regression problem, the input consists of a 256 × 256 RGB representation of the squealer-tip geometry, while the output is a dense single-channel 256 × 256 temperature field. The temperature values are normalized to the [0, 1] range during training and subsequently converted back to their physical values for post-processing. Consequently, the model predicts one temperature value at each of the 256^2 pixels instead of providing a single scalar prediction for the entire geometry.

A U-Net architecture was adopted for this purpose. Originally introduced for image segmentation, U-Net consists of an encoder that progressively reduces the spatial resolution while extracting increasingly complex features and a symmetric decoder that reconstructs the output field. Its main characteristic is the presence of *skip connections* between encoder and decoder layers at corresponding spatial resolutions. These connections concatenate high-resolution features extracted in the encoder with the upsampled features generated by the decoder, allowing local geometric information that could otherwise be lost during downsampling to be preserved. This characteristic is particularly useful in the present application, where accurately reconstructing localized temperature variations associated with the cooling-hole positions is essential. A more detailed theoretical description of U-Net architectures can be found in Siddique et al. [25].

The architecture adopted in this work is shown in Fig. 4.

The encoder progressively decreases the spatial resolution through Max Pooling while increasing the number of feature channels. The decoder then reconstructs the field through progressive upsampling and concatenation with the corresponding encoder features. Each convolutional block contains two 3 × 3 convolutional layers, each followed by a normalization layer and an activation function. The complete architecture is summarized in Table 4.

Group Normalization (GN) was used instead of Batch Normalization (BN). BN estimates a mean and variance across all samples in a batch for each channel. With the small batch sizes imposed by graphics-processing-unit (GPU) memory limitations, these statistics can become unstable and cause large loss fluctuations. GN instead divides the channels into groups and computes the normalization statistics independently for each image. It is therefore independent of the batch size and provides more stable training under the present conditions.

3.2.3. U-net masking, loss, and training

The U-Net was trained using AdamW, a variant of Adam in which weight decay is decoupled from the gradient-based parameter update. This formulation provides more stable regularization and can improve the generalization of deep convolutional networks.

The geometry images include pixels that do not belong to the physical tip surface, namely the background and the interiors of the cooling holes. A geometry-dependent binary mask M was therefore generated from every input image, as illustrated in Fig. 5. The mask equals 1 on the physical squealer-tip domain and 0 elsewhere. During training and evaluation, the pixel-wise loss and performance metrics are computed only where $M = 1$.

Operationally, the loss associated with each pixel is multiplied by its mask value. Pixels on the physical domain therefore contribute to the weight update, whereas masked-out pixels do not. Nevertheless, the network still returns a scalar value at every position of the full 256 × 256 grid. Because values at pixels for which $M = 0$ are not

Table 4
U-Net architecture used for temperature-field prediction.

Layer	Operation	Output resolution	Channels
Input	RGB image	256 × 256	3
Encoder Block 1	DoubleConv	256 × 256	64
Encoder Block 2	MaxPool + DoubleConv	128 × 128	128
Encoder Block 3	MaxPool + DoubleConv	64 × 64	256
Encoder Block 4	MaxPool + DoubleConv	32 × 32	512
Bottleneck	MaxPool + DoubleConv	16 × 16	1024
Decoder Block 1	Upsample + Skip + DoubleConv	32 × 32	512
Decoder Block 2	Upsample + Skip + DoubleConv	64 × 64	256
Decoder Block 3	Upsample + Skip + DoubleConv	128 × 128	128
Decoder Block 4	Upsample + Skip + DoubleConv	256 × 256	64
Output	1 × 1 Convolution	256 × 256	1

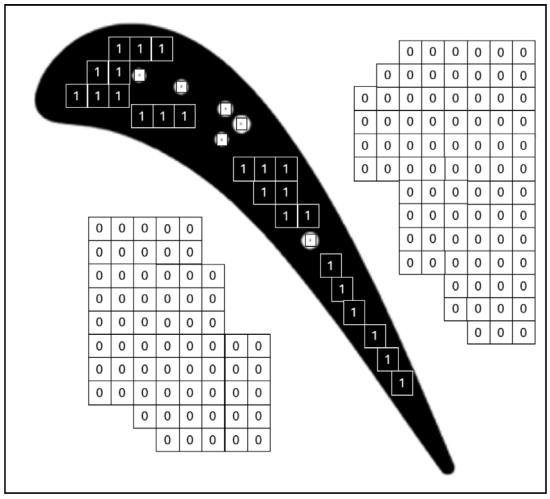


Fig. 5. Conceptual binary mask used to identify the physical squealer-tip domain and exclude external pixels.

constrained by the training loss, they may be arbitrary and have no physical meaning. They are never used when evaluating a candidate design.

The same geometry-dependent mask is regenerated and reapplied for every geometry evaluated during optimization. Background and cooling-hole pixels are therefore excluded candidate by candidate from the RMSD and all other spatial metrics. The masked tip-averaged temperature is

$$\overline{T_{\text{tip}}} = \frac{\sum_{i=1}^{256^2} M_i \hat{T}_i}{\sum_{i=1}^{256^2} M_i}, \tag{2}$$

where $M_i \in \{0, 1\}$ is the geometry-dependent mask and $\hat{T}_i$ is the predicted temperature at pixel i. Consequently, only pixels for which $M_i = 1$ contribute to averaged quantities and optimization objectives.

The U-Net loss combines a masked SmoothL1, or Huber, term with a spatial-gradient term. For the normalized temperature error e, the SmoothL1 component is

$$L_{\text{SmoothL1}} = \begin{cases} \frac{1}{2}\frac{e^2}{\delta}, & |e| < \delta, \\ |e| - \frac{1}{2}\delta, & |e| \geq \delta, \end{cases} \tag{3}$$

where δ controls the transition from quadratic to linear behavior. In this work, $\delta = 0.02$, corresponding to approximately 2% of the normalized temperature range. The gradient component penalizes discrepancies between the spatial gradients of the predicted and reference fields, encouraging physically consistent temperature distributions. The total loss is

$$L_{\text{tot}} = L_{\text{SmoothL1}} + \lambda_{\text{grad}} L_{\text{grad}}, \tag{4}$$

where λ_{grad} controls the relative weight of the gradient loss.

Table 5
Hyperparameter search ranges for the U-Net model.

Hyperparameter	Search range
Batch Size	[4–16]
Dropout Value	[0.00, 0.02, 0.05, 0.08, 0.10, 0.12, 0.15]
Activation Function	[ReLU, LeakyReLU]
Learning Rate α_{LR}	$[10^{-5.2} - 10^{-2.6}]$
Weight Decay λ_{wd}	$[10^{-7} - 10^{-3.5}]$

The coefficient of determination R^2 was also monitored during training. To limit GPU-memory usage, it was evaluated for each batch on a random subset of 20 000 pixels belonging to the masked physical domain. This provides a representative estimate of prediction accuracy at a manageable computational cost.

Temperature values were normalized between 0 and 1 during training. During post-processing, the predicted values were rescaled to the physical temperature interval used to train the model. The U-Net hyperparameters were selected through random search. Owing to the substantially higher training cost of the convolutional architecture relative to the ANNs, the search was limited to 10 configurations, each assessed using 3-fold cross-validation. The explored variables were batch size, dropout probability, activation function, learning rate, and weight decay, with learning rate and weight decay sampled logarithmically. Training was allowed to continue for a maximum of 5000 epochs, with early stopping based on validation loss and a patience of 80 epochs. The complete search space is reported in Table 5.

3.2.4. RGB-to-temperature target reconstruction

The CFD thermal contours were exported as RGB images, whereas training the U-Net requires a scalar reference temperature value at each pixel. Therefore, the RGB contour images had to be converted into scalar temperature maps using the look-up table (LUT) associated with the *rainbow uniform* CFD colormap.

A direct LUT-interpolation strategy was initially tested. The RGB channels, originally ranging from 0 to 255, were normalized between 0 and 1 and used as coordinates in a three-dimensional RGB interpolation space. The interpolator then estimated a temperature for each pixel from the LUT. This method proved inaccurate because the colors contained in the exported CFD contours did not always coincide exactly with the discrete RGB combinations stored in the LUT. Continuous interpolation in RGB space therefore introduced systematic reconstruction errors and produced temperature fields that differed significantly from the expected physical values.

To overcome this limitation, a Nearest Color Mapping (NCM) procedure was adopted. Each pixel color was associated with the closest RGB triplet available in the LUT rather than being continuously interpolated in RGB space. A KDTree was used to accelerate the nearest-neighbor search. The LUT entries are interpreted as points in three-dimensional RGB space, and the similarity between the pixel (R, G, B) and LUT entry (R_i, G_i, B_i) is measured through the Euclidean distance

$$D_{\text{RGB},i} = \sqrt{(R - R_i)^2 + (G - G_i)^2 + (B - B_i)^2}. \tag{5}$$

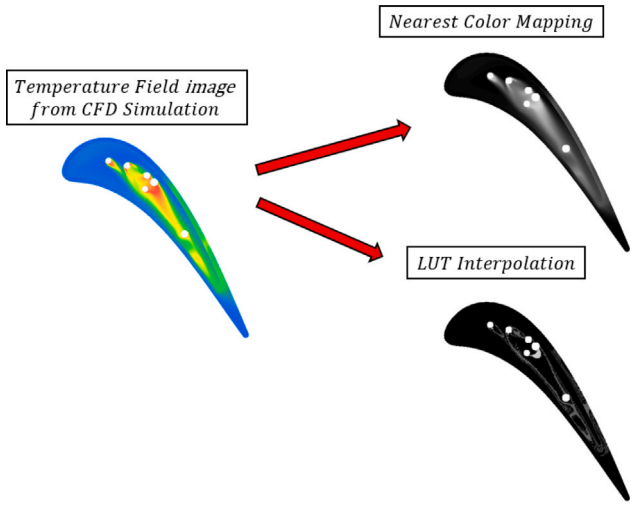


Fig. 6. Comparison of direct LUT interpolation and nearest-color mapping for recovering scalar temperature values from RGB images.

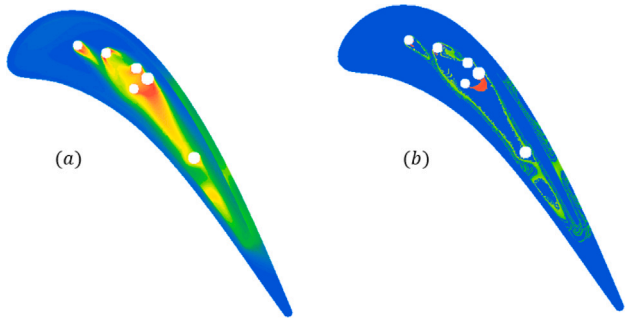


Fig. 7. Reverse LUT mapping applied to the scalar fields obtained using (a) nearest-color mapping and (b) direct LUT interpolation.

For each CFD-image pixel, the KDTree identifies the LUT entry that minimizes this distance. The corresponding LUT index is then used to recover the scalar temperature. This method remains robust when exported colors differ slightly from the tabulated RGB values and provides a consistent reconstruction of the temperature field. The resulting scalar maps constitute the reference targets used to train the U-Net.

The comparison in Fig. 6 shows that NCM reconstructs the scalar temperature distribution accurately, whereas direct LUT interpolation produces a coarser and less accurate result. Once a scalar field has been obtained or predicted, the LUT can also be applied in the reverse direction to generate an RGB contour for visualization, as shown in Fig. 7. This reverse mapping is not part of U-Net training or inference; it is used only for visual comparison of the reconstructed fields.

This RGB-to-scalar conversion inevitably introduces a quantization error. However, a direct point-by-point comparison between the original CFD temperature field and the reconstructed scalar field is not possible, since the CFD solution is defined on the computational mesh, whereas the reconstructed field is available on a fixed 256 × 256 pixel grid. Therefore, to quantify this overall error, two complementary analyses were performed. First, the $\overline{\overline{T_{tip}}}$ obtained directly from the CFD solution was compared with the one computed from the reconstructed scalar image over the entire dataset. Therefore, this comparison accounts for both the RGB quantization error and the error introduced by the image discretization of the CFD solution. The resulting mean absolute effectiveness error was found to be only 0.15 K or 0.00484 in term of cooling effectiveness, confirming that the reconstruction accuracy is fully adequate for the present application.

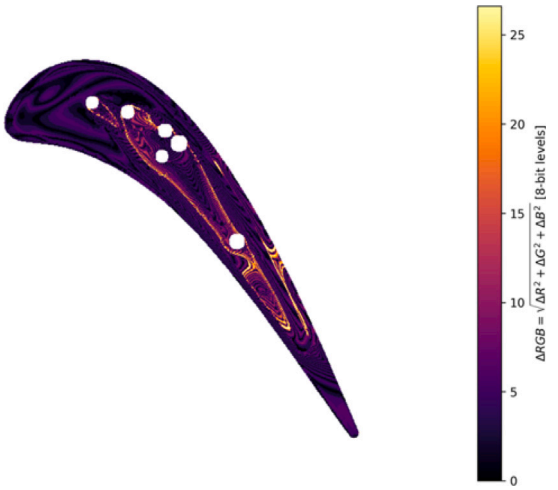


Fig. 8. Distribution of the temperature quantization error introduced by the RGB-to-scalar conversion.

As a second validation, the reconstructed scalar field was converted back into an RGB image using the inverse LUT mapping and compared with the original CFD RGB image. The pixel-wise color difference was quantified through

$$\Delta_{\mathrm{RGB}} = \sqrt{(\Delta R)^2 + (\Delta G)^2 + (\Delta B)^2}, \tag{6}$$

where ΔR, ΔG, and ΔB denote the differences in the three color channels. The average value computed over the entire dataset (considering only the physical tip domain) was found to be 4.1. An example of the resulting Δ_{RGB} distribution is reported in Fig. 8.

3.3. Optimization framework

The trained surrogate models were coupled with PSO to evaluate the four objective functions described below. All objectives are formulated for minimization. Depending on the selected function, candidate designs are evaluated either exclusively through the two scalar ANNs or through a combination of the U-Net and the coolant-mass-flow ANN.

3.3.1. Objective functions and normalization

Four objective functions were considered to demonstrate how the same surrogate framework can identify different solutions according to the selected design priority:

$$\theta_1 = -\left(\overline{\overline{\eta_{\mathrm{aw}}}}\right)_{\mathrm{norm}}, \tag{7}$$

$$\theta_2 = \left(\mathrm{RMSD}(T, T_{0,\mathrm{cool}})\right)_{\mathrm{norm}}, \tag{8}$$

$$\theta_3 = \left(\dot{m}_{\mathrm{cool}}\right)_{\mathrm{norm}} - \left(\overline{\overline{\eta_{\mathrm{aw}}}}\right)_{\mathrm{norm}}, \tag{9}$$

$$\theta_4 = \left(\dot{m}_{\mathrm{cool}}\right)_{\mathrm{norm}} + \left(\mathrm{RMSD}(T, T_{0,\mathrm{cool}})\right)_{\mathrm{norm}}. \tag{10}$$

A generic quantity z is normalized using the minimum and maximum values contained in the DoE:

$$z_{\mathrm{norm}} = \frac{z - z_{\min}}{z_{\max} - z_{\min}}, \qquad 0 \le z_{\mathrm{norm}} \le 1. \tag{11}$$

The root-mean-square deviation of the temperature field from the coolant total temperature is defined as

$$\mathrm{RMSD}(T, T_{0,\mathrm{cool}}) = \sqrt{\frac{1}{N_\Omega} \sum_{i \in \Omega} \left(T_i - T_{0,\mathrm{cool}}\right)^2}, \tag{12}$$

where Ω is the set of masked pixels belonging to the physical tip domain, N_Ω is the number of pixels in Ω, and T_i is the temperature at pixel i. Objectives θ_1 and θ_3 rely on the ANN predictions, whereas θ_2 and θ_4 require the complete U-Net field. The RMSD exploits the spatial

information reconstructed by the U-Net: it promotes more complete coolant coverage by strongly penalizing regions whose temperature remains far from $T_{0,\mathrm{cool}}$. An optimization based only on $\overline{\overline{\eta_{\mathrm{aw}}}}$ could instead favor fields containing very cold regions alternating with poorly cooled hot regions.

Under the present experimental conditions, $T_{0,\mathrm{cool}} > T_{0,\mathrm{main}}$. Maximizing $\overline{\overline{\eta_{\mathrm{aw}}}}$ therefore also identifies the configuration with the maximum tip-averaged temperature $\overline{\overline{T_{\mathrm{tip}}}}$. Because PSO is implemented as a minimization algorithm, effectiveness appears with a negative sign in θ_1 and θ_3.

3.3.2. Particle swarm optimization and constraints

Particle Swarm Optimization is a population-based optimization method inspired by collective behavior such as bird flocking and fish schooling [26]. The tentative solutions form a swarm, and each solution is represented by a particle. Particle motion depends on the best position previously found by that particle, $\mathbf{p}_{i,\mathrm{best}}$, and the best position found by the full swarm, $\mathbf{g}_{\mathrm{best}}$.

At iteration j, the velocity $\mathbf{v}_i^{(j)}$ and position $\mathbf{z}_i^{(j)}$ of particle i are updated according to

$$\mathbf{v}_i^{(j+1)} = \omega \mathbf{v}_i^{(j)} + c_1 r_1 \left(\mathbf{p}_{i,\mathrm{best}} - \mathbf{z}_i^{(j)}\right) + c_2 r_2 \left(\mathbf{g}_{\mathrm{best}} - \mathbf{z}_i^{(j)}\right), \tag{13}$$

$$\mathbf{z}_i^{(j+1)} = \mathbf{z}_i^{(j)} + \mathbf{v}_i^{(j+1)}. \tag{14}$$

Here, ω is the inertia weight, c_1 and c_2 are the cognitive and social coefficients, and r_1 and r_2 are independent random values in $[0, 1]$ that are resampled for every particle and iteration. The optimization terminates when the maximum number of iterations $J_{\max}$ is reached or when a convergence criterion is satisfied, for example

$$\left\| \mathbf{g}_{\mathrm{best}}^{(j+1)} - \mathbf{g}_{\mathrm{best}}^{(j)} \right\|_2 < \tau_{\mathrm{PSO}}. \tag{15}$$

The standard PSO procedure was modified to enforce the physical constraints of the present geometry. Active holes must remain inside the squealer cavity and must not overlap. In addition, each active hole must satisfy a prescribed positive minimum clearance from the cavity boundary and from every other active hole. First, swarm initialization uses the same conditional sampling procedure adopted for the DoE: x_i is sampled within its admissible range, and y_i is sampled between the local values of y_{upper} and y_{lower}. This directs the initial population toward feasible configurations.

Second, the minimized fitness function contains the selected objective together with a penalty term. A penalty of 5 is added whenever a hole lies outside the region bounded by y_{upper} and y_{lower} after accounting for the prescribed wall clearance, or whenever two active holes overlap or violate the prescribed hole-to-hole clearance. Feasible designs are consequently favored throughout the iterations. This formulation allows the clearance value to be adapted to manufacturing-process capabilities or to structural-design requirements without changing the general optimization procedure.

The activation variables remain continuous inside the optimizer, which can explore intermediate values between 0 and 1. Before a particle is passed to the surrogate models, each activation variable is converted to a Boolean value: values greater than 0.5 are set to 1 and values below 0.5 are set to 0. The threshold could be increased or reduced to favor fewer or more active holes. During the evolution of the swarm, the activation variables naturally tend to approach values close to 0 or 1 as information is transferred from one population to the next.

3.3.3. Artificial neural network-based optimization loop

The ANN-based workflow is illustrated in Fig. 9. The first swarm population is initialized in terms of the 44 geometric variables. For every particle, the two ANNs predict $\overline{\overline{\eta_{\mathrm{aw}}}}$ and $\dot{m}_{\mathrm{cool}}$, and the required objective function θ is then evaluated. If the iteration index j has reached the prescribed maximum $J_{\max}$, the best objective value and its corresponding geometric parameters are returned. Otherwise, PSO updates the 44 variables for the complete population and the evaluation is repeated. The execution time required for the ANN optimization framework, using a PSO with 4000 particles and 20 iterations, is approximately 2 min.

3.3.4. U-net-based optimization loop

The U-Net-based workflow is shown in Fig. 10. PSO still operates on the same 44-variable geometric representation. However, because the U-Net requires an image rather than a parameter vector, a dedicated Python script converts every candidate particle into an RGB image of the squealer geometry and its active cooling holes.

The U-Net predicts the complete normalized temperature field from this geometry image, from which RMSD$(T, T_{0,\mathrm{cool}})$ and the other masked spatial quantities are evaluated. The coolant mass flow rate is still obtained from the corresponding ANN. A CNN does not provide a specific benefit for predicting this scalar quantity and would require a substantially greater computational cost than the fully connected ANN. Conversely, the U-Net is advantageous for the temperature prediction because it retains the spatial organization of the thermal field and provides information about the local interaction of the cooling jets.

The predicted temperature values are denormalized using the minimum and maximum temperatures employed during training. When a colored contour is required for visualization, each scalar temperature can then be associated with an RGB color through direct interpolation along the LUT temperature scale. Unlike the inaccurate RGB-to-temperature interpolation tested during target generation, this reverse transformation interpolates a scalar temperature that already lies within the LUT range and is used only to display the predicted field.

For each candidate geometry, the binary mask is regenerated and reapplied before computing any spatial metric that are therefore not calculated over the complete 256 × 256 output grid, but only over pixels for which the mask equals 1. Values predicted in the background and inside cooling holes are ignored because they were not constrained during training. This masked processing recovers the same global mean-temperature quantity represented by the scalar ANN while retaining the full local field. It therefore enables more advanced objectives that promote complete thermal coverage and minimize both the mean response and the RMSD from the coolant temperature.

The execution time required by the ANN + U-Net optimization framework, using a PSO with 4000 particles and 20 iterations, is longer than that of the ANN-based framework, reaching approximately 10 h. Nevertheless, the computational cost remains extremely low, considering that performing the same 80,000 objective-function evaluations using CFD would require approximately 2.8 million minutes, corresponding to about 46,700 h, or 5.3 years of continuous computation.

4. Results

4.1. Models performances

4.1.1. Artificial neural network performances

A total of 2000 CFD simulations were performed in order to generate the dataset used for training the neural networks. Since this is a relatively small network, it was possible to train the networks using one central processing unit (CPU) core, with the entire random search training process requiring only about 24 h of computation. Instead of adopting a single multi-output network, two separate single-output networks were employed: one for the adiabatic effectiveness prediction (whose performance is reported in Tables 6 and 7) and one for the mass flow rate (whose performance is reported in Tables 8 and 9).

Both networks are able to effectively predict the output as a function of the input parameters, achieving satisfactory R^2 values considering the strong complexity of the proposed parametrization. The R^2 obtained for the mass flow rate is slightly lower than that of the temperature and shows a small degree of overfitting. For completeness,

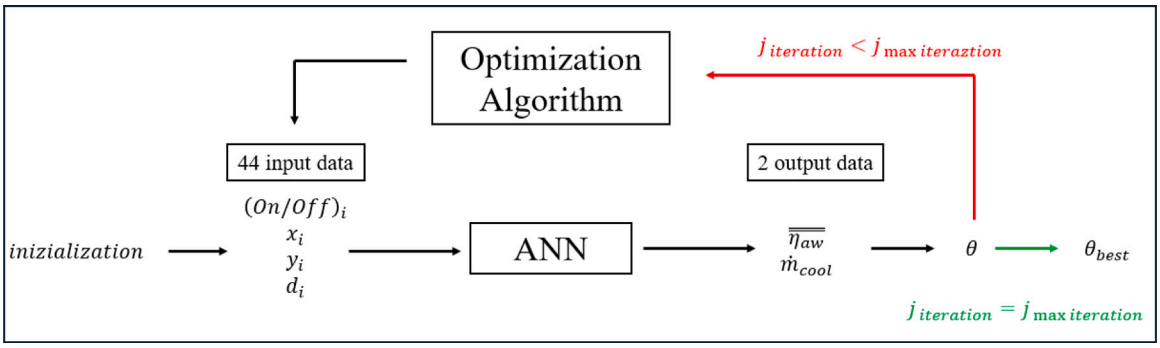


Fig. 9. ANN-based optimization workflow coupling cooling-hole parametrization, surrogate evaluation of scalar metrics, and PSO.

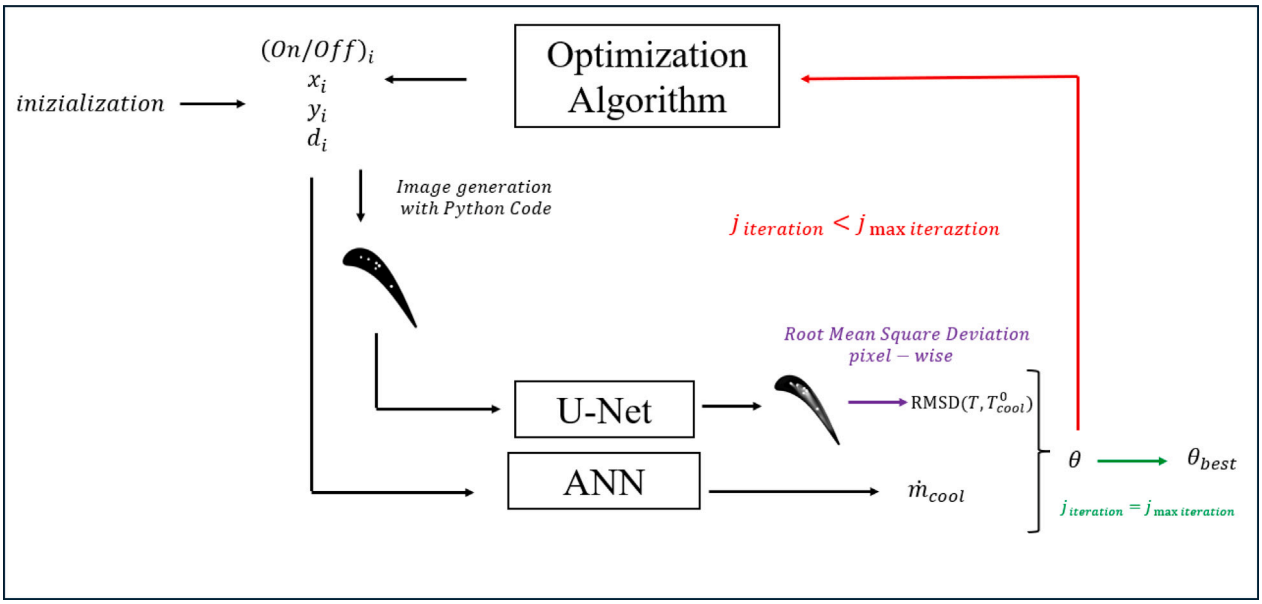


Fig. 10. Combined ANN–U-Net optimization workflow using scalar and full-field surrogate predictions within PSO.

Table 6
ANN performance for $\overline{\overline{\eta_{aw}}}$ - R^2.

R^2	Value
R^2 training	0.9467
R^2 validation	0.9181
R^2 test	0.9311

Table 7
ANN performance for $\overline{\overline{\eta_{aw}}}$ - Loss.

Loss	Value
Loss training	0.0039546
Loss validation	0.004312
Loss test	0.004689

Table 8
ANN performance for $\dot{m}_{cool}$ - R^2.

R^2	Value
R^2 training	0.9155
R^2 validation	0.8933
R^2 test	0.8732

the parity plots for training, validation, and test sets are also reported. The temperature parity plot for the test dataset is shown in Fig. 11, while the corresponding plot for the mass flow rate is reported in Fig. 12.

Table 9
ANN performance for $\dot{m}_{cool}$ - Loss.

Loss	Value
Loss training	0.0036401
Loss validation	0.0040008
Loss test	0.0053161

Moreover, the convergence plots show that for both quantities the training process reaches convergence after approximately 300 epochs, with both the loss and the R^2 stabilizing. The training is therefore automatically stopped through the early stopping criterion implemented via the patience parameter. Training and validation loss histories for the ANN are reported in Figs. 13 and 14.

Finally, the optimal hyperparameters obtained through the random search procedure are reported in Tables 10 and 11. In both cases, a relatively small neural network, composed of only two hidden layers and using the LeakyReLU activation function (negative slope = 0.01), proved to be the most effective architecture.

4.1.2. U-Net performances

The training of the neural networks was performed using the HPC@PoliTO high-performance computing infrastructure at Politecnico di Torino, equipped with NVIDIA H200 GPUs. The training of each network required approximately 5 h of computation on a single GPU. Considering the entire random search procedure and the associated

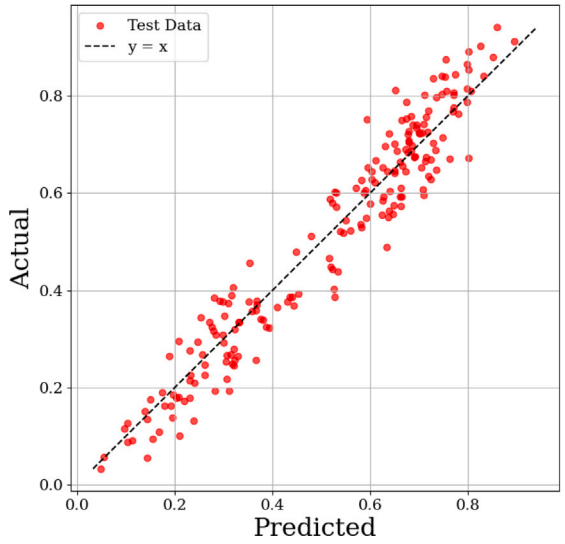


Fig. 11. Test-set parity plot for $\overline{\overline{\eta_{aw}}}$: ANN predictions versus unseen CFD targets.

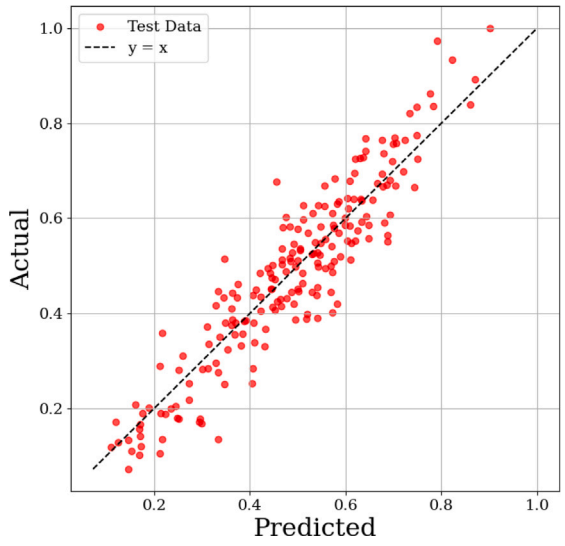


Fig. 12. Test-set parity plot for $\dot{m}_{cool}$: ANN predictions versus unseen CFD targets.

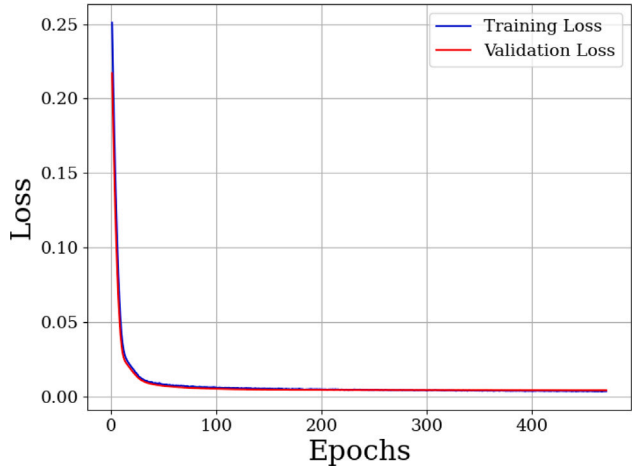


Fig. 13. Training and validation loss histories for the ANN predicting $\overline{\overline{\eta_{aw}}}$.

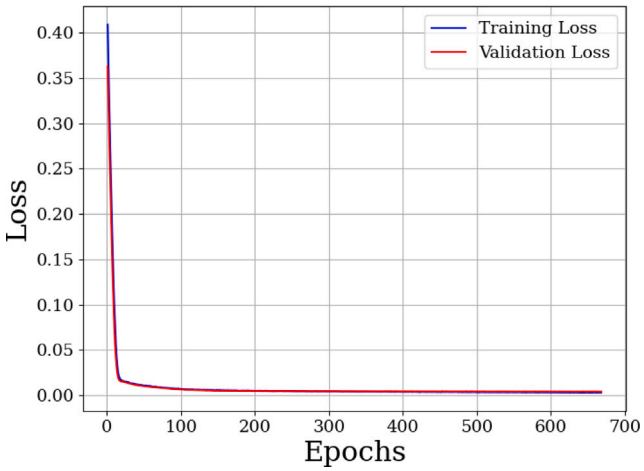


Fig. 14. Training and validation loss histories for the ANN predicting $\dot{m}_{cool}$.

Table 10
Optimal hyperparameters from random search for $\overline{\overline{\eta_{aw}}}$.

Hyperparameter	Optimal value
Best hidden layer size	[122, 11]
Best learning rate	0.000086403
Best batch size	13
Best weight decay	0.000095908
Best dropout value	0.15
Best activation function	LeakyReLU (negative slope=0.01)

Table 11
Optimal hyperparameters from random search for $\dot{m}_{cool}$.

Hyperparameter	Optimal value
Best hidden layer size	[122, 164]
Best learning rate	0.000014192
Best batch size	47
Best weight decay	0.000056442
Best dropout value	0.05
Best activation function	LeakyReLU (negative slope=0.01)

Table 12
U-Net performance.

R^2	Value
R^2 training	0.9926
R^2 validation	0.9494
R^2 test	0.9373

Table 13
U-Net performance.

Loss	Value
Loss training	0.0069146
Loss validation	0.020101
Loss test	0.020716

cross-validations, the total GPU time amounted to approximately 150 h. The performance of the U-Net in predicting the thermal field is reported in Tables 12 and 13, while the optimal hyperparameters are listed in Table 14. Among the tested configurations, the LeakyReLU activation function was found to provide the best performance. Considering the complexity of the thermal scenario, the U-Net achieves extremely good predictive capabilities. Although a slight degree of overfitting is observed, the network demonstrates a strong ability to reconstruct the scalar thermal field at the blade tip.

The convergence histories of the loss during training is shown in Fig. 15, indicating that convergence is reached only after approximately 800 epochs.

Table 14
Optimal hyperparameters from random search for U-Net.

Hyperparameter	Optimal value
Best learning rate	0.000014547
Best batch size	16
Best weight decay	0.00000022841
Best dropout value	0.15
Best activation function	LeakyReLU (negative slope=0.01)

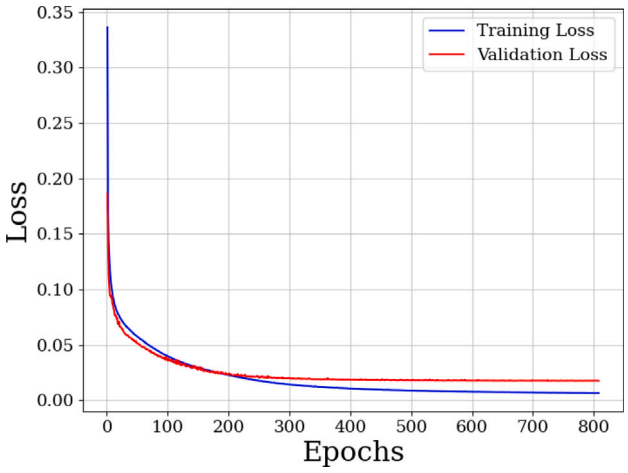


Fig. 15. Training and validation loss histories of the U-Net temperature-field model.

In this case, it is also possible, and particularly effective from a qualitative standpoint, to directly compare the temperature distribution predicted by the U-Net with the one obtained through CFD simulations. Such a comparison is reported in Fig. 16 for a geometry belonging to the test dataset (therefore not seen by the network during training). The U-Net is able to accurately reproduce the thermal field, demonstrating not only the capability of predicting the behavior of individual cooling holes, but also the complex temperature patterns arising from strong interactions among multiple cooling jets. As a result, the two fields appear almost indistinguishable to the naked eye. At the same time, once the network has been trained on the CFD-generated dataset, the thermal field can be predicted in just a few seconds, reducing the optimization time to only a few minutes instead of several hours of numerical simulations, thereby enabling an extremely fast redesign capability.

To better visualize the error distribution, an error map was generated both in terms of absolute temperature error and relative error (computed with respect to the CFD temperature at each pixel), as shown in Fig. 17. The errors are almost negligible over most of the domain, except in some localized hot regions downstream of the cooling holes, particularly the first two, where the coolant coverage is higher (in this case the coolant temperature is higher than the mainstream temperature). In these regions the maximum error reaches approximately 5.92 K. It should be noted that the U-Net prediction is inherently scalar, while the maps are displayed in RGB format only for visualization purposes and to allow a direct image-based comparison through the use of the LUT.

4.1.3. Statistical analysis of surrogate-model error

An additional independent database of 1100 CFD DPs was generated specifically to assess the surrogate-model error. The cases were sampled separately for each number of active cooling holes, from 1 to 11, using a Sobol distribution, with exactly 100 DPs generated for each hole-count class. This strategy provides the same statistical weight to all values of the number of active holes and avoids biasing the analysis toward more frequently represented configurations.

Table 15
Statistical indicators of the normalized absolute validation error evaluated on the 1100 additional CFD data points.

Model	Mean [%]	Median [%]	P_{95} [%]	$2\sigma_0$ [%]	Max [%]
$\dot{m}_{\mathrm{cool}}$-ANN	2.600	2.091	6.528	6.627	12.259
$\overline{\overline{T_{\mathrm{tip}}}}$-ANN	1.403	1.264	3.197	3.427	5.463
$\overline{\overline{T_{\mathrm{tip}}}}$-CNN	1.239	1.034	3.130	3.195	5.995

For a generic predicted scalar quantity y, the normalized absolute error is defined as:

$$\varepsilon_y = 100\,\frac{|y_{\mathrm{ML}} - y_{\mathrm{CFD}}|}{y_{\max} - y_{\min}}, \tag{16}$$

where the CFD solution is used as reference. Normalization by the complete excursion of the corresponding quantity provides an error measure directly related to the physically relevant prediction range, rather than to the magnitude of the individual reference value. The adopted ranges are $\Delta\dot{m} = 6.9292 \times 10^{-3}$ kg/s for coolant mass flow rate and $\Delta T = 31$ K for temperature.

A direct comparison between the ANN and CNN thermal predictions requires both models to be evaluated using the same scalar quantity. Therefore, the temperature field reconstructed by the U-Net is spatially averaged over the valid squealer-tip mask, and the resulting $\overline{\overline{T_{\mathrm{tip}}}}$ is compared with the corresponding CFD value. The same quantity is directly predicted by the ANN. In this way, the reported thermal errors provide a consistent comparison between the two approaches, although the U-Net additionally reconstructs the complete spatial temperature distribution.

The resulting error distributions are reported in Fig. 18, while the main statistical indicators are summarized in Table 15. The coolant mass-flow ANN exhibits the largest prediction uncertainty, with a mean normalized error of 2.60% and a 95th percentile of 6.53%. The thermal models show lower errors, with mean values of 1.40% for the ANN and 1.24% for the CNN. Their upper error tails are also comparable, with 95th-percentile values of approximately 3.20% and 3.13%, respectively.

The slightly lower mean and median errors obtained with the CNN are particularly relevant considering that its prediction involves the additional image-based representation and reconstruction of the temperature field. Nevertheless, both thermal approaches provide similar overall levels of accuracy when evaluated on $\overline{\overline{T_{\mathrm{tip}}}}$. More importantly for their subsequent use within the optimization framework, the characteristic mean surrogate errors remain lower than the appreciable performance differences observed among the designs generated by the different optimization objectives, which are presented in the following section. Therefore, the models retain sufficient accuracy to discriminate the main design trends. Conversely, very small differences between closely performing candidates should not be interpreted solely from surrogate predictions, motivating the independent CFD re-evaluation of the final optimized configurations.

To investigate whether the prediction accuracy is related to specific characteristics of the cooling-hole arrangement, four geometrical descriptors were considered: the number of active cooling holes N_{act}, the mean diameter of the active holes, their total active area, and the mean pairwise distance between active-hole centroids. The total active-hole area combines the effects of the number and diameter of the holes and is defined as:

$$A_{\mathrm{act}} = \sum_{i=1}^{11} (\mathrm{On/Off})_i \frac{\pi d_i^2}{4}, \tag{17}$$

where $(\mathrm{On/Off})_i$ is equal to one for an active hole and zero otherwise. Fig. 19 reports the mean normalized error within each geometrical bin for the three surrogate outputs, while the shaded bands indicate one standard deviation around the corresponding mean. The number of

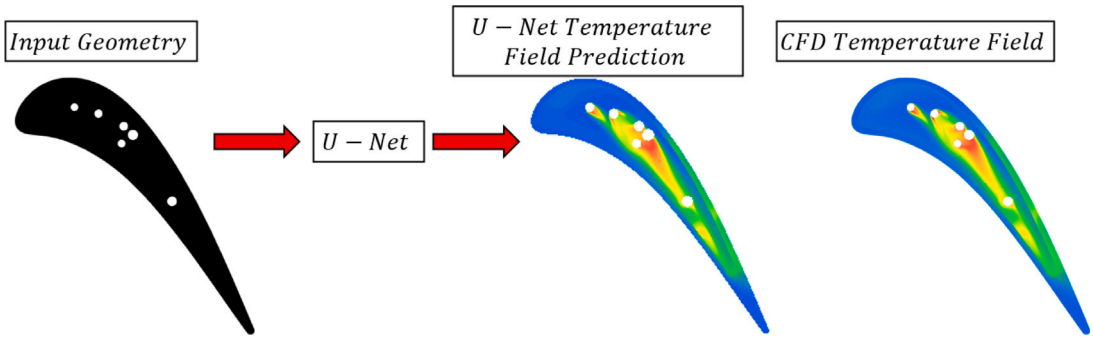


Fig. 16. Representative comparison between the U-Net predicted temperature field and the corresponding CFD reference.

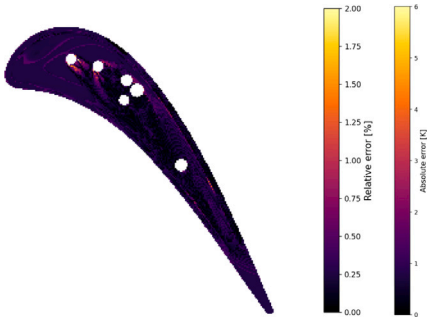


Fig. 17. Spatial distributions of the absolute and normalized percentage error between the U-Net prediction and CFD reference.

observations in each bin is also reported to distinguish well-populated regions from bins for which the local statistics are less representative.

For the coolant mass-flow ANN, the average error shows a moderate increase with both the number and mean diameter of the active holes. The dependence becomes clearer when these two effects are combined through A_{act}, with configurations characterized by a larger overall discharge area generally associated with higher prediction errors. The last bins of some descriptors contain comparatively few configurations and are therefore less statistically representative.

The thermal models exhibit a less systematic dependence on the individual geometrical descriptors. Neither N_{act} nor the mean hole diameter alone identifies a clear monotonic trend common to both ANN and CNN, which is expected given the more complex interaction between hole arrangement, coolant jets, and the resulting tip-temperature distribution. A moderate increase in error is nevertheless recovered for larger values of A_{act}, suggesting that configurations with a high overall coolant-discharge area are also somewhat more challenging for the thermal surrogate models. In contrast, the mean pairwise distance between hole centroids does not show a consistent relationship with prediction accuracy. Overall, among the considered scalar descriptors, the total active-hole area provides the clearest indication of the geometrical conditions associated with increased surrogate error.

A complementary analysis was then performed to investigate how the configurations associated with the largest prediction errors are distributed across the same geometrical descriptors. Rather than considering the magnitude of the mean error within each bin, cases satisfying $\epsilon > 2\sigma_0$ were classified as statistical outliers. Fig. 20 therefore reports the relative frequency of outliers in each geometrical bin, defined as the number of outlier cases divided by the total number of configurations belonging to that bin. Using a relative frequency instead of the absolute number of outliers is necessary because the geometrical bins do not contain the same number of observations.

The outlier-frequency analysis supports the trends identified from the mean-error distributions while providing additional information on where the largest prediction deviations are concentrated. In particular, the mass-flow model shows an increasing concentration of outliers toward the upper portion of the active-area range. A similar but weaker tendency can be observed for the thermal predictions. The highest-frequency values occurring in sparsely populated bins should be interpreted cautiously, since a small number of additional outliers can produce relatively large variations in the corresponding frequency. Nevertheless, the concentration of high-error cases toward larger values of A_{act} is more consistent than the local variations observed for the other geometrical descriptors. No comparable systematic dependence is identified with the mean centroid-to-centroid distance.

4.2. Baseline flow field

The thermal distribution at the blade tip for the baseline configuration, featuring 11 cooling holes with a diameter of 1.27 mm distributed along the camber line, is shown in Fig. 21, while Fig. 22 reports the corresponding streamlines.

The flow field of the GE-E3 turbine tip in a linear cascade has been extensively investigated in the literature and is thoroughly described in [27,28]. The solution obtained in the present work is fully consistent with the results reported in those studies. As shown in Figure Fig. 22, the simulation successfully captures all the main flow structures characterizing this configuration, including the cavity flow, which develops inside the cavity after the separation induced by the step at the leading edge, the pass-over flow that bypasses the cavity without recirculating inside it, the leakage flow on the suction side generated by the pressure difference between the pressure side and the suction side, and the coolant flow injected through the cooling holes.

A strong interaction is observed particularly between the coolant flow and the cavity flow, and this interaction represents the main mechanism governing the resulting thermal field. In this configuration, the coolant jet becomes trapped inside the cavity, where it is pressed against the cavity floor and transported toward the pressure side due to the cavity flow motion. On the one hand, this behavior promotes good coolant attachment to the wall. On the other hand, as shown in Fig. 21, it leaves more than half of the blade tip surface (the region from the camber line toward the suction side) essentially uncooled. As a result, the cooling coverage is highly non-uniform, leading to an average temperature of $\overline{\overline{T_{tip}}} = 299.462$ K, an $\overline{\overline{\eta_{aw}}} = 0.0794$ with a coolant mass flow rate of $\dot{m}_{cool} = 0.001429$ kg/s.

Approximately at mid-axial chord, the coolant experiences a lift-off, moving toward the suction side, crossing the squealer gap and eventually leaving the cavity where it mixes with the leakage flow.

4.3. Constraint-violation statistic in the optimization

Before discussing the optimization results, the penalty strategy adopted to enforce geometric feasibility is briefly assessed. Since the

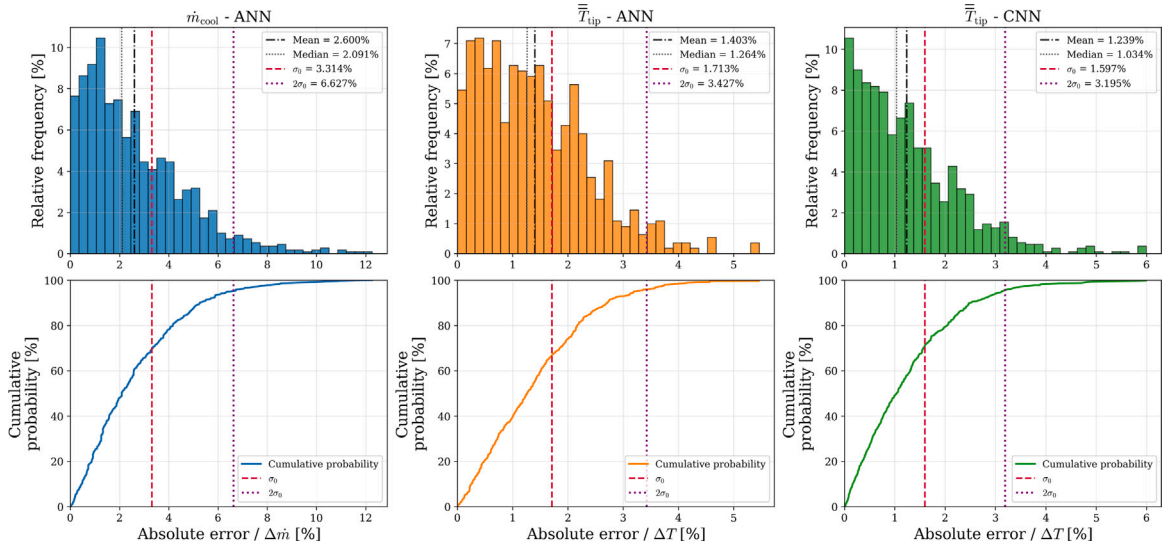


Fig. 18. Relative-frequency distributions and empirical cumulative distributions of the normalized absolute prediction error for the coolant mass-flow ANN, the area-averaged tip-temperature ANN, and the U-Net/CNN. The vertical lines identify the mean, median, $\sigma_0 = \sqrt{\text{mean}(\varepsilon^2)}$, and $2\sigma_0$.

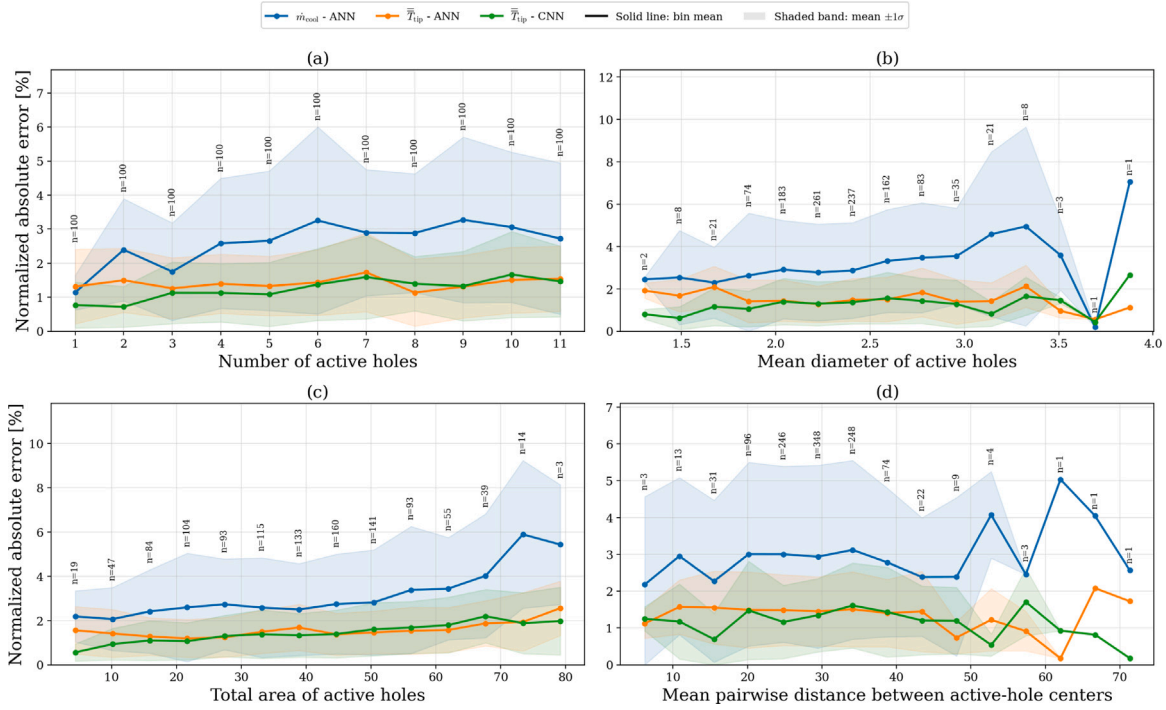


Fig. 19. Aggregated comparison of the normalized error trends as a function of: (a) number of active holes, (b) mean diameter of active holes, (c) total active-hole area, and (d) mean pairwise distance between active-hole centroids.

objective functions are normalized, their theoretical maximum excursion is 1 for the single-objective formulations and 2 for the multi-objective formulations. A penalty of 5 is therefore assigned whenever a cooling hole exits the admissible squealer region or violates the prescribed hole-to-hole or hole-to-wall clearance, ensuring that infeasible individuals cannot outperform any feasible solution. This step-like penalty is particularly suited to the evolutionary nature of PSO, since infeasible particles are progressively discarded as the swarm converges toward feasible high-performing regions of the design space.

The robustness of this approach was verified through additional PSO runs using penalty values of 1, 2, 3, 4, 5, 10, 20, and 50 and prescribed offsets of 1, 2, and 4. As shown in Fig. 23, even a penalty of 1 is sufficient to drive the fraction of invalid particles toward zero, indicating that the exact penalty magnitude has little influence on the final convergence and mainly ensures exclusion of infeasible candidates from the best solutions. The offset primarily affects the number of initially invalid configurations: smaller values reduce early constraint violations and may accelerate convergence, whereas larger clearances increase the restricted region and may initially produce more invalid individuals. Its selection should therefore be based primarily on structural and manufacturing requirements, particularly the minimum material ligament required between adjacent holes and between holes and the squealer boundaries. For all tested combinations, using 4000 particles per iteration and 20 iterations, the number of constraint-violating particles converged to zero, confirming the robustness of the adopted feasibility-enforcement strategy.

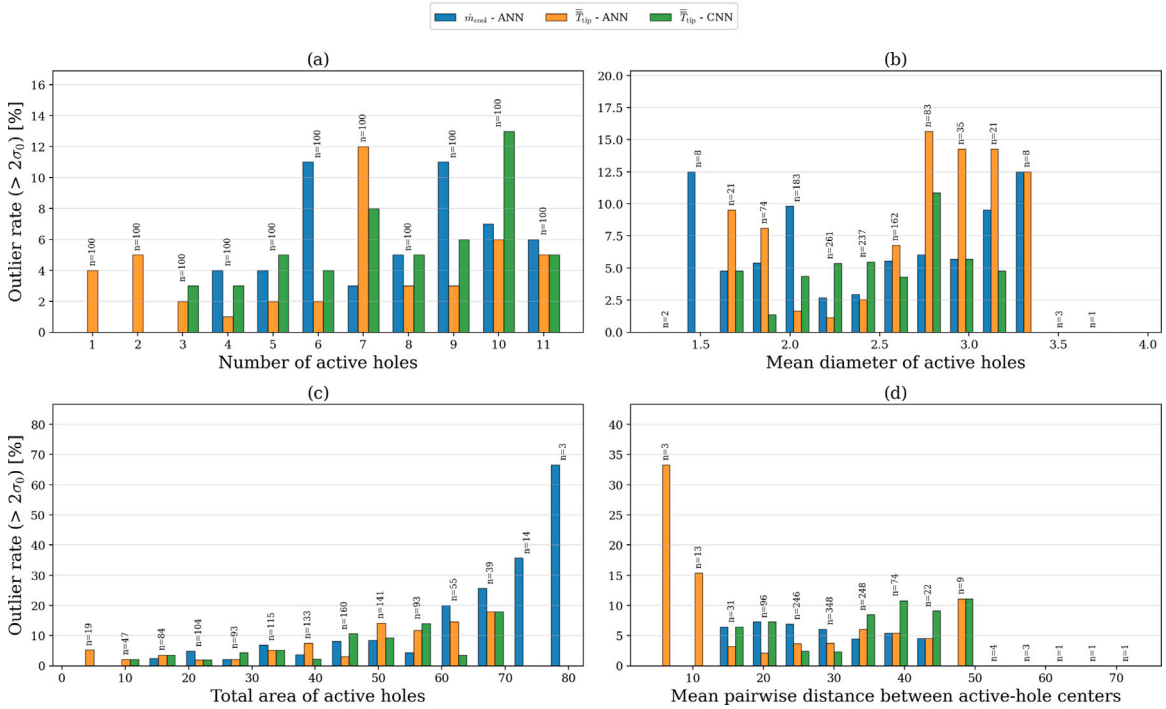


Fig. 20. Aggregated comparison of the relative frequency of outliers ($\varepsilon > 2\sigma_0$) as a function of: (a) number of active holes, (b) mean diameter of active holes, (c) total active-hole area, and (d) mean pairwise distance between active-hole centroids.

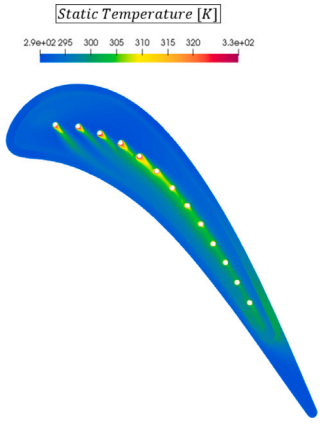


Fig. 21. CFD temperature field over the squealer tip for the baseline cooling-hole configuration.

4.4. Optimization results

The optimal temperature fields at the blade tip for each of the four optimized configurations ((a), (b), (c) and (d) based on their respective optimization functions θ_1, θ_2, θ_3, and θ_4) are shown in Fig. 24.

Once the PSO procedure had converged, the four optimized geometries were re-evaluated using independent CFD simulations. This additional step provides a direct a posteriori verification of the designs identified by the surrogate models and allows the optimized performance to be assessed using the same high-fidelity methodology adopted for the original database. The CFD values are therefore distinguished from the corresponding ANN/U-Net predictions in Table 16. Overall, the discrepancies between the surrogate predictions and the corresponding CFD results remain limited and consistent with the error levels identified in the previous validation analysis. Slightly larger deviations are observed for the first two optimized configurations, which are characterized by larger active cooling areas. In particular, the normalized errors (using the same normalization ranges adopted in the surrogate error analysis) for θ_1 are +5.10% for the coolant mass flow rate and +2.90% for the adiabatic effectiveness, while for θ_2 they are −2.85% and +2.30%, respectively. The errors remain considerably smaller for θ_3 and θ_4, below 0.75% in magnitude for both quantities. Importantly, these discrepancies do not alter the relative ranking of the optimized solutions or the main conclusions drawn from the optimization.

The first configuration (a) aims solely at maximizing the tip adiabatic effectiveness, completely disregarding the coolant mass flow rate. For this reason, the solution activates all 11 cooling holes, with 8 out of 11 reaching the maximum allowed diameter. Hole 11 and hole 5 remain at intermediate diameters of 2.369 mm and 2.2 mm, respectively, likely due to their terminal position along the camber line. Increasing their diameter further could cause the holes to exceed the lower or upper geometric boundaries, resulting in a penalty in the objective function. This configuration allows a value of $\overline{\overline{\eta_{aw}}} = 0.396$ to be achieved, at the cost of a coolant mass flow rate of $\dot{m}_{cool} = 0.008835\,kg/s$, which is approximately 6.18 times larger than the baseline value.

From the temperature contour it can be observed that the portion of the tip beyond approximately 50% of the axial chord is extremely well covered by the coolant layer, significantly increasing the average tip temperature (and therefore the film cooling effectiveness in this case). However, despite the excellent cooling coverage in the central region of the blade tip, the area closer to the trailing edge remains less protected, with entire regions of the blade tip remaining uncooled and fully exposed to the mainstream flow.

To address this limitation, configuration (b) can be considered. In this case the optimization does not only aim at maximizing $\overline{\overline{\eta_{aw}}}$, but also at minimizing the presence of uncooled regions. This is achieved through an objective function that includes the pixel-wise squared difference between the local temperature and the coolant temperature. This formulation strongly penalizes regions characterized by very low temperatures. As a result, this approach drastically reduces the presence of uncovered areas, distributing the cooling holes also toward regions that were previously unprotected. This leads to a slight decrease of $\overline{\overline{\eta_{aw}}}$ to 0.360, but at the same time the coolant mass flow rate is reduced to

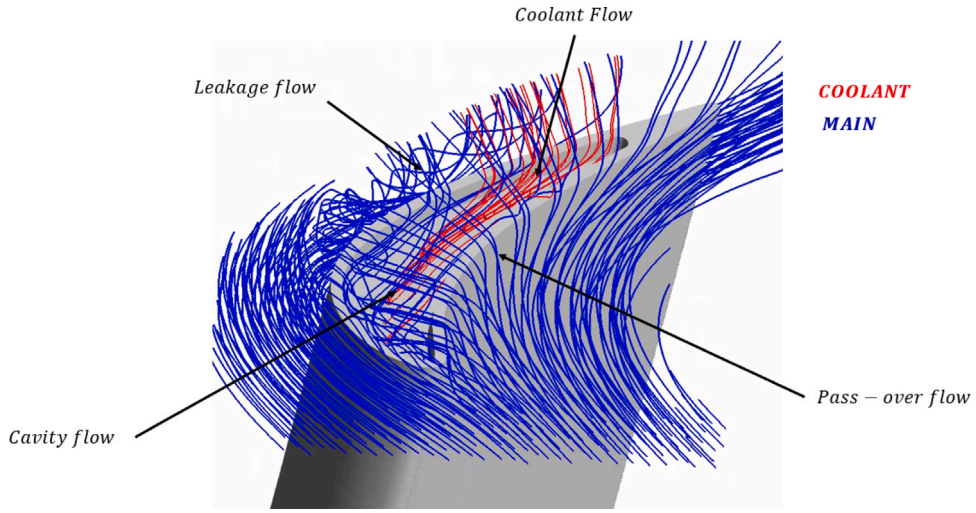


Fig. 22. CFD streamlines for the baseline configuration, showing coolant, cavity-flow, and mainstream interaction over the squealer tip.

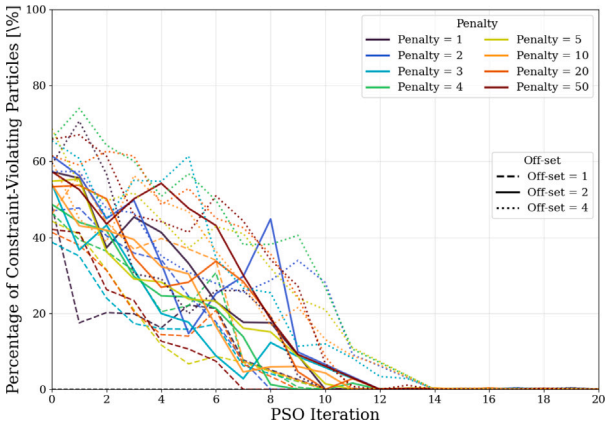


Fig. 23. Convergence of the number of particles with geometric constraint violation for different penalty values and prescribed offset values in the PSO runs.

$\dot{m}_{\text{cool}} = 0.006943\,\text{kg/s}$ (approximately 21.42% lower than configuration (a)). Furthermore, the metric $\text{RMSD}(T, T_{0,\text{cool}})$ decreases to 29.581 K compared to 29.859 K of configuration (a) and 31.496 K of the baseline configuration.

Despite this different optimization strategy, a region that is not perfectly cooled can still be observed in the upper-left area of the blade tip. The reason for this can be understood by examining Fig. 25, where it can be seen that this region coincides with a high-pressure area generated by the separation and reattachment of the flow caused by the step at the leading edge of the squealer tip. Due to this local high-pressure region, the coolant cannot effectively exit from the nearby cooling hole. Future optimization strategies could therefore focus on modifying either the hole geometry or the squealer shape to improve cooling in this area.

This type of extreme optimization, strongly biased toward temperature rather than coolant mass flow rate, is mainly intended to demonstrate the flexibility of the proposed neural-network-based framework in identifying design solutions that may differ significantly depending on the chosen design objectives.

The third configuration (c) instead aims to minimize $\dot{m}_{\text{cool}}$ while maximizing $\overline{\overline{\eta_{\text{aw}}}}$. The resulting solution includes 10 film cooling holes with diameters all close to the minimum value of 1 mm, strongly favoring mass flow reduction over $\overline{\overline{\eta_{\text{aw}}}}$. In fact, the resulting coolant mass flow rate is only $\dot{m}_{\text{cool}} = 0.000653\,\text{kg/s}$, corresponding to 45.7% of the baseline value. This reduction results from two concurrent mechanisms. The smaller hole diameters directly decrease the total coolant-discharge area, while holes located within high-pressure regions of the tip experience a larger external pressure over the tip surface that totally blocks coolant ejection. Therefore, under the prescribed plenum conditions, the coolant mass flow rate depends not only on the number and diameter of the holes, but also on their position relative to the local tip-pressure field. This leads to a slight decrease in $\overline{\overline{\eta_{\text{aw}}}}$ to 0.0705 compared to 0.0794 of the baseline configuration, which can still be considered acceptable given the significant reduction in coolant mass flow rate. However, this solution still exhibits the same limitation observed in configuration (b), namely the presence of fully uncooled regions.

Configuration (d) addresses this issue by optimizing the coolant mass flow rate together with $\text{RMSD}(T, T_{0,\text{cool}})$, instead of $\overline{\overline{\eta_{\text{aw}}}}$. This final solution includes 11 cooling holes, with 10 diameters close to the minimum value and one larger central hole with a diameter of 2.666 mm, arranged in such a way as to cover the regions that previously remained uncooled. In this configuration, $\overline{\overline{\eta_{\text{aw}}}}$ increases to 0.101, while the coolant mass flow rate rises to $\dot{m}_{\text{cool}} = 0.001153\,\text{kg/s}$. Nevertheless, this value remains approximately 19.3% lower than the baseline mass flow rate, and the metric $\text{RMSD}(T, T_{0,\text{cool}})$ decreases from 31.73 K to 31.28 K. From an engineering perspective, this can therefore be considered the best solution among the four, as it is the only configuration providing an improvement in $\overline{\overline{\eta_{\text{aw}}}}$, $\dot{m}_{\text{cool}}$, and $\text{RMSD}(T, T_{0,\text{cool}})$ simultaneously with respect to the baseline case. A summary of these results is reported in Table 16.

By analyzing these solutions, it can also be observed that even when the coolant mass flow rate is included among the optimization parameters, the algorithm tends to favor configurations with many small holes rather than solutions with fewer but larger holes. This allows the limited coolant mass flow to be distributed more effectively across the blade tip. From an engineering-design perspective, this result indicates that, for a constrained coolant budget, distributing the available flow through a larger number of small-diameter holes can provide more uniform spatial coverage than concentrating it in a few large jets. However, hole placement cannot be selected from geometrical coverage alone: a hole located in a high-pressure region may discharge little coolant despite being positioned in an otherwise poorly cooled area. Consequently, the local pressure field should be considered together with the diameter and spatial distribution of the holes when defining a practical cooling layout. As previously observed, Fig. 25 also reveals the presence of the same high-pressure region, which appears even more pronounced and extended in these configurations. This occurs because the larger holes present in configurations (a) and (b) tend to create a shielding effect, making it more difficult for the mainstream flow to

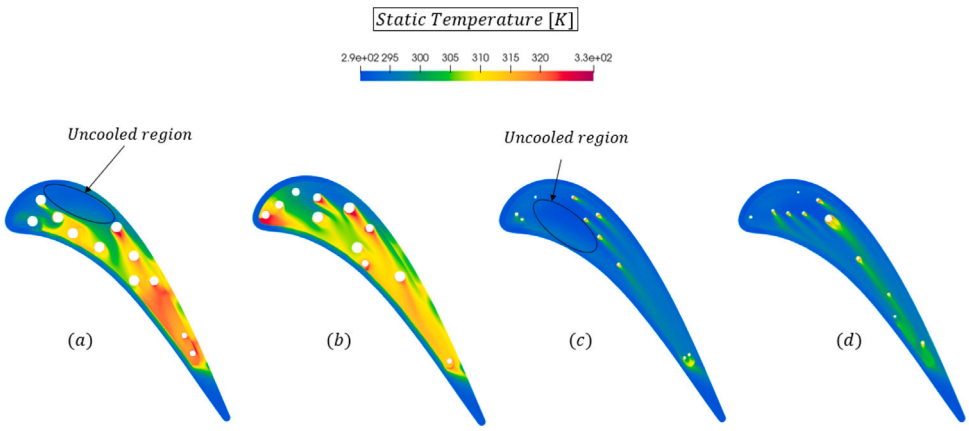


Fig. 24. CFD tip-temperature fields for the optimized configurations: (a) θ_1, (b) θ_2, (c) θ_3, and (d) θ_4.

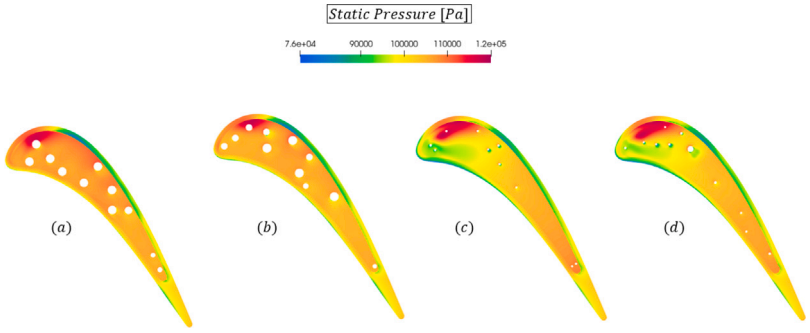


Fig. 25. CFD tip-pressure fields for the optimized configurations: (a) θ_1, (b) θ_2, (c) θ_3, and (d) θ_4.

Table 16
Performance metrics for the baseline and optimized configurations, comparing independent CFD verification with the corresponding surrogate predictions used during PSO.

Case	Objective definition	$\dot{m}_{\text{cool}}$ [kg/s]		$\overline{\overline{\eta}}_{\text{aw}}$		RMSD$(T, T_{0,\text{cool}})$ [K]	
		CFD	ANN ($\epsilon_{\dot{m}}$)	CFD	ANN ($\epsilon_{\overline{\overline{\eta}}_{\text{aw}}}$)	CFD	U-Net
Baseline		0.001429	–	0.0794	–	31.496	–
θ_1	$-\left(\overline{\overline{\eta}}_{\text{aw}}\right)_{\text{norm}}$	0.008835	0.0091884 (+5.10%)	0.396	0.425 (+2.90%)	29.859	29.807
θ_2	$\left(\text{RMSD}(T, T_{0,\text{cool}})\right)_{\text{norm}}$	0.006943	0.0067455 (−2.85%)	0.360	0.383 (+2.30%)	29.581	29.533
θ_3	$\left(\dot{m}_{\text{cool}}\right)_{\text{norm}} - \left(\overline{\overline{\eta}}_{\text{aw}}\right)_{\text{norm}}$	0.000653	0.000604 (−0.71%)	0.0705	0.0780 (+0.75%)	31.730	31.714
θ_4	$\left(\dot{m}_{\text{cool}}\right)_{\text{norm}} + \left(\text{RMSD}(T, T_{0,\text{cool}})\right)_{\text{norm}}$	0.001153	0.001105 (−0.69%)	0.101	0.0955 (−0.55%)	31.283	31.293

enter the cavity. This phenomenon can be observed more clearly in Fig. 26, which reports the streamlines for all four optimized configurations.

For the first two configurations (a) and (b), the cavity is observed to be almost entirely dominated by the coolant flow. This leads to the disappearance of the cavity flow structure that was observed in the baseline case, leaving only the presence of the pass-over flow. Only a small portion of the mainstream flow is able to enter the cavity, strongly disturbing the cooling holes located near the leading edge. This interaction reduces the cooling effectiveness of those holes in that region, which tend to rapidly experience lift-off, overcoming the squealer step and mixing with the leakage flow.

For configurations (c) and (d), a similar behavior is observed, but only for the holes located extremely close to the leading edge. For the remaining holes, the same phenomenon observed in the baseline case occurs, where the cavity flow appears to have a beneficial effect by keeping the coolant attached to the wall for a large portion of the blade tip. As a consequence, the coolant crosses the squealer step and mixes with the leakage flow only toward the end of the tip region.

This behavior is also well reflected in the optimized configurations, which show a higher density of cooling holes toward the leading edge. This region is more difficult to cool due to the presence of a high-pressure zone and the strong interaction with the cavity flow. Conversely, fewer holes are required to effectively cover regions located at larger axial coordinates.

These results provide three direct indications. First, optimizing only the area-averaged effectiveness can lead to excessive coolant consumption and can still leave localized uncooled regions; therefore, a full-field quantity such as RMSD$(T, T_{0,\text{cool}})$ should be included when spatially uniform thermal protection is required. Second, the combined mass-flow and full-field objective used for configuration (d) represents a more practical compromise between thermal protection and compressor-bleed demand than the effectiveness-only solutions. Third, the persistent leading-edge limitation suggests that simply adding cooling holes is not necessarily effective where the local pressure opposes coolant discharge; in that region, modifying the hole geometry or the squealer shape may be more beneficial than increasing the supplied coolant.

5. Conclusions

This work presented a machine-learning-based framework for the optimization of a squealer tip cooling system of the $GE - E^3$ first-stage

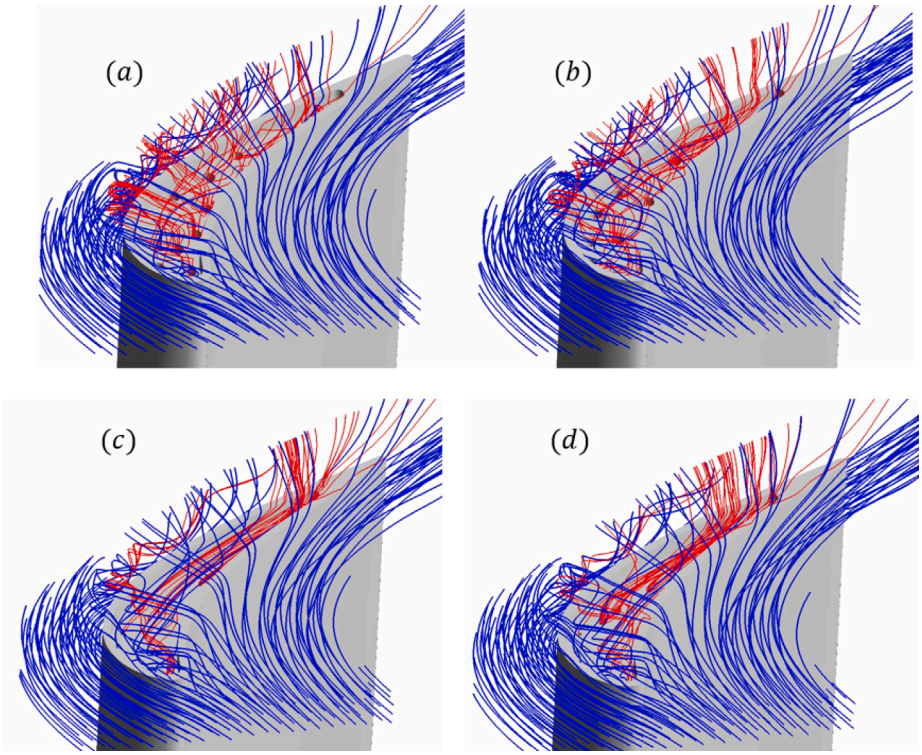


Fig. 26. CFD streamlines for the optimized configurations: (a) θ_1, (b) θ_2, (c) θ_3, and (d) θ_4.

high-pressure turbine blade. The methodology combines Computational Fluid Dynamics simulations, Artificial Neural Networks, and U-Net within a Particle Swarm Optimization loop to efficiently explore a high-dimensional design space defined by the number, diameter, and position of the cooling holes. A dataset composed of 2000 Computational Fluid Dynamics simulations was generated to train the surrogate models. Once trained, the models allow the evaluation of new geometrical configurations in a few seconds, compared to several hours required by a full Computational Fluid Dynamics computation, enabling a rapid and extensive exploration of the design space.

The main outcomes of the present study can be summarized as follows:

- The Artificial Neural Network surrogate models demonstrated good predictive capabilities for both scalar quantities involved in the optimization. The network predicting the area-averaged tip adiabatic effectiveness achieved an R^2 value of 0.923 on the test dataset, while the prediction of the coolant mass flow rate reached an R^2 of 0.873. These values are obtained despite the relatively high dimensionality of the design space (44 geometric parameters), confirming the capability of the Artificial Neural Network models to capture the nonlinear relationship between the cooling hole configuration and the thermal response of the blade tip.
- The convolutional U-Net model proved to be highly effective in reconstructing the spatial temperature distribution over the squealer tip. The model achieved an R^2 of 0.937 on the test dataset, with maximum local prediction errors of approximately 5.92 K. This result demonstrates that the convolutional architecture is able not only to predict the global thermal level but also to correctly capture the local interaction between multiple cooling jets, preserving the spatial structure of the thermal field.
- The use of the trained surrogate models within the Particle Swarm Optimization framework allowed the identification of significantly different optimal solutions depending on the selected objective function. Effectiveness-oriented optimizations were able to increase the average tip adiabatic effectiveness up to $\overline{\overline{\eta_{aw}}} = 0.396$, demonstrating the capability of the algorithm to strongly enhance coolant coverage when coolant consumption is not constrained.
- When the coolant mass flow rate is included in the objective function, the optimization converges toward configurations characterized by significantly lower coolant consumption. In particular, one optimized configuration achieved a coolant mass flow rate of $\dot{m}_{cool} = 0.000653$ kg/s, corresponding to approximately 45.7% of the baseline mass flow rate, while maintaining an average tip adiabatic effectiveness of $\overline{\overline{\eta_{aw}}} = 0.0705$. This result highlights the capability of the optimization framework to identify solutions with a substantial reduction in coolant usage while maintaining acceptable thermal protection.
- The most balanced configuration was obtained by simultaneously minimizing the coolant mass flow rate and the root mean square deviation of the temperature field from the coolant temperature. This solution achieved $\overline{\overline{\eta_{aw}}} = 0.101$ with a coolant mass flow rate of $\dot{m}_{cool} = 0.001153$ kg/s, corresponding to a reduction of approximately 19% compared to the baseline case. At the same time, the temperature root mean square deviation was reduced from 31.496 K to 31.283 K, indicating that the tip temperature field remains closer to the coolant temperature overall.
- The optimization results also revealed a clear design trend: the algorithm consistently favors configurations characterized by a larger number of small-diameter holes rather than fewer large holes. This configuration allows the limited coolant mass flow rate to be distributed more effectively across the blade tip, improving the global coverage of the cooling film. The corresponding mass-flow reduction is governed both by the smaller total discharge area and by the local back pressure imposed by high-pressure tip regions, which can partially inhibit coolant ejection.

The results presented in this study are limited to a single squealer-tip geometry, a fixed cooling-hole configuration, and a specific set of boundary conditions. Despite the improvements achieved by the

optimized configurations, the analysis highlighted the presence of a region near the leading edge of the squealer tip that remains difficult to cool effectively. This behavior is associated with the local high-pressure region generated by the separation and reattachment of the flow at the squealer step, which limits the ability of the coolant to exit from the nearby holes.

Future work should therefore investigate possible modifications of the squealer geometry to reduce the pressure peak in the leading-edge region and improve coolant coverage. In particular, variations in the squealer rim geometry or height could modify the local pressure distribution inside the cavity and promote better coolant attachment. It would also be valuable to investigate how the optimal cooling-hole configuration changes under different boundary conditions, thereby assessing the robustness of the optimized design and identifying configurations capable of maintaining high cooling performance over a wider range of operating conditions.

Furthermore, the present analysis was conducted under linear cascade conditions, and additional simulations including blade rotation effects could provide further insight into the interaction between leakage flow, cavity flow, and coolant injection. Finally, the proposed optimization framework could be extended to include additional objectives, such as aerodynamic losses or tip leakage reduction, enabling a more comprehensive multi-physics optimization of the squealer-tip design.

CRediT authorship contribution statement

Dario Antonio Orlando: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Simone Salvadori:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Daniela Anna Misul:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Computational resources were provided by hpc@polito, which is a project of Academic Computing within the Department of Control and Computer Engineering hosted and managed by the Cybersecurity and IT for Research Division (ISIAD) at Politecnico di Torino (http://hpc.polito.it).

Data availability

Data will be made available on request.

References

[1] Bunker R. Axial turbine blade tips: Function, design, and durability. J Propuls Power - J Propul Power 2006;22:271–85. http://dx.doi.org/10.2514/1.11818.
[2] Bunker R. A review of turbine blade tip heat transfer. Ann New York Acad Sci 2001;934:64–79. http://dx.doi.org/10.1111/j.1749-6632.2001.tb05843.x.
[3] Salvadori S, Insinna M, Martelli F. Unsteady flows and component interaction in turbomachinery. Int J Turbomach Propuls Power 2024;9(2):15. http://dx.doi.org/10.3390/ijtpp9020015.
[4] Zhang G, Zhu R, Xie G, Li S, Sundén B. Optimization of cooling structures in gas turbines: A review. Chin J Aeronaut 2022;35(6):18–46. http://dx.doi.org/10.1016/j.cja.2021.08.029.
[5] Che J, You R, Xu Y, Li H. Study on the effects of dimple density and dimple diameter on heat transfer and flow resistance in rotating dimpled channels. Appl Therm Eng 2025;264:125444. http://dx.doi.org/10.1016/j.applthermaleng.2025.125444.
[6] You R, Liu Y, Qi L, Che J, Xu Y, Li H. Coriolis-induced asymmetry in drag mechanisms of rotating ribbed channels. Phys Fluids 2026;38(6):066101. http://dx.doi.org/10.1063/5.0332855.
[7] De Maesschalck C, Andreoli V, Braun J, Paniagua G, Cummings W, Langford L, Bloxham M. Aerothermal optimization of fully cooled turbine blade tips. 2018, http://dx.doi.org/10.1115/GT2018-75099.
[8] De Maesschalck C, Andreoli V, Paniagua G, Gillen T, Barker B. Aerothermal optimization of turbine squealer tip geometries with arbitrary cooling injection. J Turbomach 2021;143:1–12. http://dx.doi.org/10.1115/1.4051268.
[9] Duan P, He L. Optimization of turbine cascade squealer tip cooling design by combining shaping and flow injection. J Turbomach 2021;143:1–12. http://dx.doi.org/10.1115/1.4051204.
[10] Guo J, Song L, Tao Z, Li J. Cooling layout optimization for a turbine blade squealer tip with the application of oval holes. Int Commun Heat Mass Transfer 2024;159:108323. http://dx.doi.org/10.1016/j.icheatmasstransfer.2024.108323.
[11] Vincekovic L, Shahpar S. Multidisciplinary optimisation of high pressure turbine blade tip with turbine inlet capacity and stage reaction constraints. 2021, http://dx.doi.org/10.2514/6.2021-3071.
[12] Li S, Du X, Gao Y, Fan A, Zhao Y, Du H. Enhanced energy storage performance by stabilizing antiferroelectric phase through high-entropy engineering of NaNbO3. J Mater Sci, Mater Electron 2026;37(12):888. http://dx.doi.org/10.1007/s10854-026-17285-x.
[13] Gao Y, Du X, Li Z, Wang L, Tian B, Tian R, Du H. Enhanced VOCs degradation performance of TiO2-OV/NH2-MIL-125 by synergistic modulation of oxygen vacancy and heterojunction. Colloids Surfaces A: Physicochem Eng Asp 2026;741:140318. http://dx.doi.org/10.1016/j.colsurfa.2026.140318.
[14] Wang L, Lu J, Li Z, Wu L, Gao Y, Fu Y, Ran J, Du H. Lattice matching of Zn2V2O7/NaV8O20 cathode enhances built-in electric field for durable zinc-ion batteries. J Electroanal Chem 2026;1015:120228. http://dx.doi.org/10.1016/j.jelechem.2026.120228.
[15] Lei M, Zhang M, Wang K. Research on bidding optimization strategy for virtual power plants with wind–solar–storage systems based on IGDT–DRO. Electr Eng 2026;108(3):208. http://dx.doi.org/10.1007/s00202-026-03544-x.
[16] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707. http://dx.doi.org/10.1016/j.jcp.2018.10.045.
[17] Huang X, Han Z, Deng D, Li J, Qi S, Han S, Zhang R, Zhou L, Li L, Zhang H, Du X. Predicting distribution of adiabatic film cooling effectiveness using generative adversarial networks with mist/air mixture as coolant. Phys Fluids 2026;38(6):063312. http://dx.doi.org/10.1063/5.0316011.
[18] Liu Z, Han S, Zhang R. Parameter coupling mechanism and film cooling effectiveness prediction of serrated-V-shaped composite structure of turbine blade based on GRNN model. Int J Therm Sci 2026;228:110973. http://dx.doi.org/10.1016/j.ijthermalsci.2026.110973.
[19] Tao Y, Liu Z, Zhang W, Feng Z. Optimization of film cooling and aerodynamic performance for asymmetric cavity bottom blade tip based on generative adversarial network surrogate model. Aerosp Sci Technol 2026;168:110932. http://dx.doi.org/10.1016/j.ast.2025.110932.
[20] Labrini G, Nastasi R, Salvadori S, Misul DA. Optimization of an high-pressure turbine vane for pressure gain combustion cycles through convolutional neural network. Energy AI 2026;24:100729. http://dx.doi.org/10.1016/j.egyai.2026.100729.
[21] Yıldız F, Alpman E, Kavurmacıoğlu L, Camci C. Aerothermal optimization of film cooling hole locations on the squealer tip of an HP turbine blade. J Therm Sci Eng Appl 2024;16(5):051001. http://dx.doi.org/10.1115/1.4064431.
[22] Kwak JS, Han J-C. Heat transfer coefficients and film-cooling effectiveness on a gas turbine blade tip. J Heat Transf 2003;125:494. http://dx.doi.org/10.1115/1.1565096.
[23] Azad G, Han J-C, Boyle R. Heat transfer and flow on the squealer tip of a gas turbine blade. J Turbomachinery-Trans Asme - J Turbomach-T ASME 2000;122. http://dx.doi.org/10.1115/1.1311284.
[24] Purwono I, Ma'arif A, Rahmaniar W, Imam H, Fathurrahman HIK, Frisky A, Haq QMU. Understanding of convolutional neural network (CNN): A review. Int J Robot Control Syst 2023;2:739–48. http://dx.doi.org/10.31763/ijrcs.v2i4.888.
[25] Siddique N, Sidike P, Elkin C, Devabhaktuni V. U-net and its variants for medical image segmentation: theory and applications. 2020, http://dx.doi.org/10.48550/arXiv.2011.01118.
[26] Ampellio E, Bertini F, Ferrero A, Larocca F, Vassio L. Turbomachinery design by a swarm-based optimization method coupled with a CFD solver. Adv Aircr Spacecr Sci 2016;3:149–70. http://dx.doi.org/10.12989/aas.2016.3.2.149.
[27] Yan X, Huang Y, He K. Investigations into heat transfer and film cooling effect on a squealer-winglet blade tip. Int J Heat Mass Transfer 2017;115:955–78. http://dx.doi.org/10.1016/j.ijheatmasstransfer.2017.08.090.
[28] Huang M, Li Z, Li J. Investigations on the aerothermal performance of the turbine blade winglet squealer tip within an uncertainty framework. Aerosp Sci Technol 2022;123:107506. http://dx.doi.org/10.1016/j.ast.2022.107506.