

# Trasformazione digitale nella produzione dell'acciaio: un sistema di monitoraggio e sicurezza basato su Intelligenza Artificiale per le operazioni dei forni elettrici

V. Colla, M. Waseem Akram, A. Siddique, M. Vannucci, G. Bavestrelli, R. Girelli

Il progetto iSteel-Expert introduce un sistema innovativo di monitoraggio e sicurezza basato su intelligenza artificiale, specificamente progettato per le operazioni dei forni elettrici ad arco nelle acciaierie. Il sistema integra una rete di sensori e telecamere con una piattaforma di raccolta dati in tempo reale, utilizzando modelli avanzati di deep learning per automatizzare il rilevamento e la classificazione precisa di eventi critici e condizioni di sicurezza nell'ambiente del forno. Questo approccio innovativo migliora significativamente sia gli standard di sicurezza operativa che l'efficienza complessiva del processo. Il framework di monitoraggio si concentra su diversi aspetti operativi chiave: il rilevamento delle persone nell'area del forno per valutare l'esposizione umana e la conformità degli operatori alle procedure stabilite, il controllo dello stato della porta di scorifica per verificare la conformità dei processi di colata, il monitoraggio continuo del movimento del braccio degli elettrodi attraverso il rilevamento della posizione delle pinze per identificare anomalie che potrebbero richiedere correzioni o manutenzione, e il rilevamento di vapori, fumi e fiamme per garantire la sicurezza ambientale e operativa.

Un dataset completo di circa 5000 immagini è stato raccolto per l'addestramento e la validazione dei modelli. Il sistema ha dimostrato ottime prestazioni, raggiungendo un'accuratezza superiore al 90% per il rilevamento dell'inclinazione del forno e del fumo, con elevati valori di precisione, recall e F1-score nei compiti di object detection.

**PAROLE CHIAVE:** FORNO AD ARCO ELETTRICO; SICUREZZA; IMAGE PROCESSING; DEEP LEARNING; INTELLIGENZA ARTIFICIALE.

## INTRODUZIONE

L'industria siderurgica ha implementato diverse misure di sicurezza, che comprendono programmi di formazione, strategie di manutenzione all'avanguardia e una rigorosa gestione operativa [1]. Nonostante i progressi ottenuti, sussiste l'esigenza di perfezionare i metodi di identificazione dei pericoli attraverso la transizione da procedure manuali a processi interamente automatizzati. Questa trasformazione è vitale per accrescere l'efficacia degli interventi di sicurezza all'interno degli impianti siderurgici. I sistemi automatizzati sono in grado di individuare precocemente i segnali di potenziali incidenti tramite il monitoraggio continuo dell'ambiente operativo. L'Intelligenza Artificiale (AI) può essere sfruttata per l'analisi dei dati, la rilevazione di anomalie e le attività di sorveglianza in tali contesti industriali. Inoltre, i sistemi di analisi basati su AI contribuiscono a una riduzione dei costi operativi tramite l'automatizzazione di compiti precedentemen-

**Valentina Colla, Muhammad Waseem Akram, Arslan Siddique, Marco Vannucci**

Scuola Superiore Sant'Anna Istituto TeCIP, Pisa, Italy

**Giovanni Bavestrelli, Renato Girelli**

TENOVA S.p.a., Castellanza, Varese, Italy

te eseguiti manualmente. Il Deep Learning si è affermato come una tecnologia estremamente potente per svariate applicazioni di AI, come il rilevamento e il tracciamento di oggetti. In questo studio, impieghiamo il deep learning con l'obiettivo di elevare i livelli di sicurezza e l'efficienza operativa negli impianti di produzione dell'acciaio.

Le operazioni condotte nei forni elettrici ad arco (EAF), sebbene essenziali per il riciclo dell'acciaio, espongono i lavoratori a seri pericoli per la sicurezza derivanti da elementi quali gli elettrodi in grafite, gli intensi archi elettrici, l'alta tensione e le emissioni di gas nocivi [2]. Per affrontare questa problematica, ci siamo concentrati sulla minimizzazione dell'esposizione umana negli ambienti EAF e sull'abilitazione di un sistema per la identificazione rapida delle situazioni pericolose attraverso l'applicazione della visione artificiale. Questo lavoro ha portato allo sviluppo di EAFvision, un sistema automatizzato per la sorveglianza di sicurezza in quasi real-time. Un sistema integrato di telecamere, microfoni e sensori aggiuntivi è stato installato in un sito EAF operativo per acquisire dataset specifici sulle criticità di sicurezza. Grazie all'addestramento e alla validazione di algoritmi di visione artificiale su questi dati, siamo ora in grado di rilevare tempestivamente i segnali premonitori di incidenti, identificando fumo, fiamme e condizioni anomale del forno. Inoltre, EAFvision monitora la presenza dei lavoratori per dare la possibilità di interrompere automaticamente le operazioni a rischio non appena viene rilevato personale all'interno delle aree critiche.

I metodi utilizzati per il rilevamento degli oggetti possono essere classificati in due categorie principali: i rilevatori a due stadi e i rilevatori a stadio singolo. I rilevatori a due stadi, come Faster R-CNN [3], Mask R-CNN [4] e Cascade R-CNN [5], procedono generando inizialmente delle proposte di regioni in cui è probabile la presenza degli oggetti, per poi classificarle e perfezionarle in una fase successiva. A differenza di questi, i rilevatori a stadio singolo eseguono entrambe le operazioni contemporaneamente, risultando quindi più rapidi ed efficaci per le applicazioni che richiedono l'elaborazione in tempo reale. Tra i numerosi rilevatori a stadio singolo, i modelli della famiglia YOLO (You Only Look Once) [6-9] hanno ottenuto notevole attenzione per le loro capacità di rilevamento in tempo reale. Abbiamo condotto esperimenti con alcuni dei modelli YOLO più recenti, per individuare le migliori

prestazioni e ottenere una comparazione dettagliata delle loro performance sui nostri dataset. Oltre ai modelli YOLO, sono stati sviluppati anche altri tipi di rilevatori a stadio singolo [10]. Negli ultimi anni, i transformer sono emersi come una potente alternativa alle tradizionali reti convoluzionali per l'elaborazione dei dati visivi, fra cui il modello DEtection TRansformer (DETR) [11]. Il modello RT-DETR [12] affronta le sfide relative alla velocità e all'efficienza introducendo meccanismi di attenzione ottimizzati e una strategia di addestramento migliorata per accelerare il processo di rilevamento pur mantenendo un'elevata accuratezza. Dato il successo di RT-DETR, ne abbiamo eseguito l'addestramento e il testing con il nostro dataset, realizzando un confronto approfondito delle prestazioni con altri rilevatori sia a stadio singolo che a due stadi.

Questo lavoro di ricerca colma un'importante lacuna nell'applicazione delle tecniche di visione artificiale per il miglioramento della sicurezza e dell'efficienza operativa nell'ambito della produzione di acciaio, con un focus specifico sugli ambienti EAF. Contrariamente agli studi precedenti che spesso si concentravano su compiti singoli o un numero ristretto di modelli, la nostra ricerca valuta un'ampia gamma di modelli di rilevamento oggetti all'avanguardia, incluse le versioni recenti di YOLO e i rilevatori basati su transformer, per compiti di sicurezza essenziali quali la rilevazione del personale, il monitoraggio delle pinze per gli elettrodi e l'identificazione del fumo. Questo studio fornisce un'analisi comparativa approfondita dell'accuratezza dei modelli, della velocità di inferenza e della fattibilità della loro implementazione su dispositivi edge. Evidenziando le performance di modelli con requisiti computazionali ridotti come YOLOv9s, che combinano alta precisione con bassa latenza, offriamo linee guida concrete per lo sviluppo di sistemi di monitoraggio in quasi real-time e che ottimizzino la sicurezza e i flussi di produzione nei contesti industriali.

## MATERIALE E METODI

Il presente lavoro di ricerca analizza l'impiego delle tecniche di deep learning e di visione artificiale con l'obiettivo di migliorare i livelli di sicurezza e ottimizzare l'efficienza operativa nel settore siderurgico, con particolare attenzione al contesto altamente critico degli EAF. A tal fine, è stata condotta un'attività sistematica di ottimizzazione

e valutazione di diversi algoritmi di object detection allo stato dell'arte, con l'obiettivo di analizzarne le prestazioni e l'efficacia in una serie di compiti di rilevamento ritenuti critici. Tali attività sono state progettate specificamente per rafforzare la sicurezza e migliorare l'efficienza dei processi operativi negli ambienti EAF.

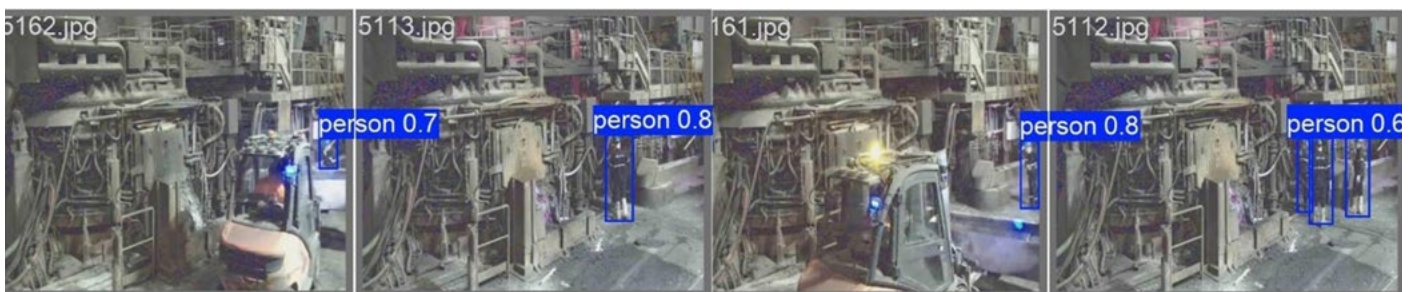
### Task selezionati

Sono state individuate tre attività fondamentali per rafforzare le misure di sicurezza e garantire un monitoraggio efficace dei processi operativi più critici all'interno dell'ambiente EAF. La loro selezione, tra diverse possibili alternative, è stata guidata dalla rilevanza e dalla criticità rispetto alle principali problematiche di sicurezza di un impianto EAF in esercizio. Le attività considerate sono le seguenti:

1. Rilevamento del personale: Uno degli obiettivi principali nello sviluppo di sistemi di visione artificiale per l'ambiente EAF è l'incremento della sicurezza dei lavoratori. A tal fine, i sistemi proposti sono in grado di attivare automaticamente un allarme sonoro o visivo quando un operatore viene rilevato all'interno di un'area operativa critica, segnalando tempestivamente una potenziale situazione di rischio. Questo compito di rilevamento, focalizzato su una singola classe identificata come *persona*, consente il monitoraggio continuo della presenza del personale. La Figura 1 mostra alcuni esempi rappresentativi dell'attività di rilevamento del personale in impianti EAF.
2. Monitoraggio delle pinze dell'elettrodo: Questa attività è finalizzata all'individuazione e al tracciamento della posizione delle pinze degli elettrodi, al fine di

garantire che il loro movimento all'interno del forno elettrico ad arco avvenga correttamente. Deviazioni o anomalie nel comportamento degli elettrodi possono infatti indicare guasti meccanici o malfunzionamenti operativi. Il sistema è progettato per segnalare tali irregolarità agli operatori in modo tempestivo, consentendo un intervento immediato. Anche in questo caso si tratta di un compito di rilevamento a singola classe, etichettata come *pinza*. La figura 2 riporta esempi del rilevamento delle pinze.

3. Identificazione del fumo nelle vicinanze degli elettrodi: La terza attività prevede il monitoraggio continuo delle emissioni nelle aree adiacenti agli elettrodi, con l'obiettivo di rilevare la presenza di fumo e individuare eventuali concentrazioni anomale. Un'identificazione precoce di tali condizioni è cruciale per riconoscere irregolarità o criticità nel processo EAF. Al superamento di soglie considerate pericolose, il sistema genera avvisi che permettono agli operatori di intervenire rapidamente, riducendo il rischio di escalation. Questa funzione è formulata come un compito di classificazione binaria, volto a distinguere tra la presenza e l'assenza di fumo. La Figura 3 presenta alcuni esempi visivi dell'attività di rilevamento del fumo. A differenza dei task precedenti, questo è formulato come classificazione binaria a livello di immagine, per la quale è stata sfruttata la modalità di classificazione nativa del framework Ultralytics, che supporta YOLO anche come classificatore puro (senza bounding box). Questa scelta garantisce uniformità tecnologica con gli altri task e semplifica il deployment sul dispositivo edge.



**Fig.1** - Esempio di immagini inerenti al task di rilevamento delle persone in prossimità dell'EAF. La bounding box identifica l'oggetto della ricerca e fornisce una stima della probabilità di accuratezza della classificazione (nel range [0;1]) / Example images related to the task of detecting personnel in proximity to the EAF. The bounding box identifies the target object and provides an estimate of the classification confidence score (in the range [0, 1]).



**Fig.2** - Esempio di immagini inerenti al task di rilevamento della pinza degli elettrodi / *Example images related to the task of detecting the electrode clamp.*



**Fig.3** - Esempio di immagini inerenti il task di classificazione riguardo alla presenza di fumo in prossimità degli elettrodi. La label nell'immagine indica l'esito della classificazione / *Example images related to the classification task concerning the presence of smoke in proximity to the electrodes. The label in the image indicates the classification outcome.*

## DESCRIZIONE DEL DATASET

Per lo svolgimento di questo studio è stato utilizzato un dataset acquisito direttamente da un impianto EAF operativo. All'interno dell'impianto sono stati installati sistemi di monitoraggio comprendenti telecamere, microfoni e sensori aggiuntivi, impiegati per il controllo di diversi parametri di processo. " Il dataset è stato acquisito nell'arco di circa un anno di funzionamento ordinario dell'impianto, coprendo in modo naturale le diverse condizioni operative tipicamente incontrate durante la produzione. Le immagini sono state annotate da un esperto operante all'interno dell'impianto siderurgico, con competenza diretta sui processi EAF; le annotazioni sono state successivamente revisionate da ulteriori esperti di impianto al fine di garantirne la qualità e la coerenza. Le condizioni ambientali riflettono fedelmente quelle di un impianto industriale reale: il posizionamento delle telecamere è stato determinato dai vincoli fisici e operativi del sito, e il dataset include inevitabilmente condizioni di variabilità ambientale quali presenza di polvere, vapore e variazioni di illuminazione, che possono influenzare la qualità delle

immagini. Queste condizioni, pur rappresentando una sfida per i modelli, costituiscono al contempo una garanzia di rappresentatività rispetto agli scenari operativi reali. Il numero di immagini disponibili per ciascuna attività è riportato in Tabella 1.

Al fine di garantire un addestramento accurato e una valutazione oggettiva e priva di bias dei modelli, il dataset è stato suddiviso casualmente, per ciascun task, in tre sottoinsiemi distinti: training, validazione e test. Le annotazioni finali così ottenute sono state quindi elaborate e rese disponibili per le successive fasi di addestramento e valutazione dei modelli.

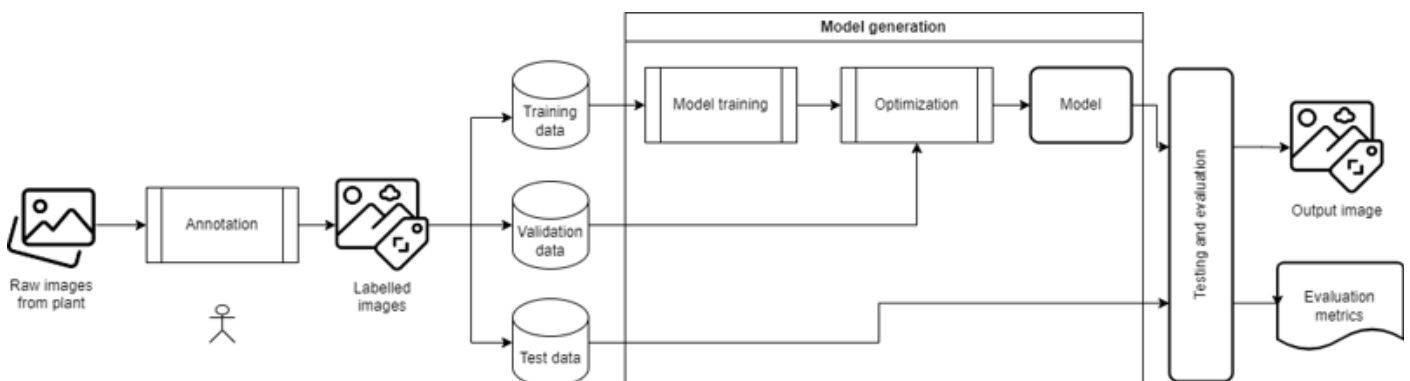
**Tab.1** - Numero di immagini disponibili per i diversi task / Number of available images for the different tasks.

Task di rilevamento	Training samples	Validation samples	Test samples
Pinze elettrodi	788	112	114
Persone	2427	305	307
Fumo	429	91	94

### PIPELINE DI ELABORAZIONE

La figura 4 mostra la pipeline per l'addestramento e il testing dei modelli di visione artificiale utilizzati in questo lavoro. Il processo inizia con le immagini raccolte dalle telecamere presenti sull'impianto, che vengono annotate utilizzando uno strumento open-source. Il dataset annotato viene suddiviso in set di addestramento, validazione e test, con le suddivisioni di addestramento e validazione

utilizzate per l'addestramento del modello. Durante l'addestramento, un algoritmo di apprendimento identifica pattern nei dati, mentre il set di validazione aiuta a ottimizzare gli iperparametri e ad evitare l'overfitting, garantendo la generalizzazione su nuovi dati. Dopo l'addestramento, il modello viene sottoposto a ottimizzazione dei parametri per ridurre il tempo di inferenza.

**Fig.4** - Pipeline per l'addestramento e la valutazione dei modelli / Pipeline for the training and evaluation of the models.

### CAMPAGNA SPERIMENTALE

Per l'attività sperimentale, si è proceduto con l'addestramento di diverse varianti del modello YOLO (dalle versioni 8 alla 11) e di RT-DETR, sfruttando il framework *Ultralytics*. Contemporaneamente, sono stati addestrati RetinaNet, Mask R-CNN, Faster R-CNN e Cascade R-CNN, utilizzando il framework *Detectron2* [13]. Per il task riguardante la detezione del fumo si è utilizzata la modalità classify anziché detect resa disponibile dal framework utilizzato. Tali modelli sono stati poi sottoposti a valutazione su un dataset personalizzato.

I modelli YOLO sono stati addestrati per un totale di 20 epochs, adottando una learning rate iniziale di 0.01. Per gli altri modelli, che comprendono Faster R-CNN, Mask R-CNN, RetinaNet e Cascade R-CNN, è stato impiegato l'ottimizzatore AdamW [14] con una learning rate fissata a 0.002. L'addestramento di questi modelli è stato esteso per 1500 epochs, garantendo un numero sufficiente di cicli per apprendere le strutture complesse presenti nei dati. Durante la fase di addestramento sono state applicate tecniche di aumento dei dati (data augmentation), come ResizeShortestEdge, per accrescere la robustezza

dei modelli. È importante sottolineare che il numero di epochs non è direttamente confrontabile tra i due framework: i modelli YOLO, addestrati tramite Ultralytics, beneficiano di tecniche di data augmentation integrate e di una strategia di apprendimento ottimizzata che favorisce una convergenza rapida. L'analisi delle curve di loss ha confermato che 20 epochs erano sufficienti per raggiungere la convergenza sui dataset utilizzati. Per i modelli Detectron2 (Faster R-CNN, Mask R-CNN, RetinaNet e Cascade R-CNN), il numero maggiore di epochs riflette le differenze intrinseche nell'architettura di training del framework. Gli esperimenti sono stati eseguiti su un sistema di calcolo ad alte prestazioni, operativo su ambiente Ubuntu e dotato di una GPU NVIDIA GeForce RTX 3090 e 24 GB di RAM. Ciò ha assicurato un addestramento efficiente e la possibilità di effettuare l'inferenza in tempo reale per le attività di deep learning. Per lo sviluppo dei modelli YOLO sono stati utilizzati Python 3.10.5 e PyTorch 2.5.1, mentre l'implementazione dei modelli Detectron2 è avvenuta con Python 3.9.21 e PyTorch 2.5.1 con supporto CUDA, necessario per gestire le esigenze computazionali di dataset estesi e architetture complesse.

In conformità con il protocollo di valutazione adottato da Lan in [2] e che è ormai uno standard, si è fatto ricorso alla mean Average Precision (mAP), alla precisione, alla *recall* e all'*F1* score come metriche di valutazione principali [15].

## RISULTATI

In questa sezione, vengono presentati i risultati conseguiti per i tre task che sono stati trattati. L'analisi condotta considera sia l'accuratezza delle prestazioni ottenute dai modelli, sia la loro effettiva applicabilità all'interno di un contesto industriale. Nel quadro generale delle valutazioni, si è osservato che i modelli appartenenti alla famiglia YOLO hanno fornito performance di alta qualità se confrontati con altre architetture di *deep learning*. Di conseguenza, si è proceduto alla valutazione sia della rapidità di esecuzione dei modelli sia della loro accuratezza.

### Test dei modelli in quasi real-time

La tabella 2 riporta i tempi necessari per l'inferenza, misurati direttamente sul sistema destinato all'utilizzo dei modelli nell'ambiente industriale.

A seguito di un processo di ottimizzazione, il modello

YOLOv9s è stato convertito nel formato TensorFlow Lite (TFLite), rendendolo idoneo per un'implementazione efficiente su dispositivi edge. Per consentire la previsione in quasi real-time, è stata creata un'applicazione in Python che sfrutta RabbitMQ RPC ed è stata inclusa in un container Docker per facilitarne la scalabilità. Il modello è stato quindi installato su un dispositivo Tenova EDGE con un processore Intel® Celeron® J3455, 8 GB di RAM e sistema operativo AdvLinuxTU, configurando l'ambiente ideale per il monitoraggio della sicurezza. I tempi di inferenza riportati in Tabella 2 sono stati misurati direttamente sul dispositivo Tenova EDGE (Intel® Celeron® J3455, 8 GB di RAM), con i modelli convertiti in formato TensorFlow Lite (TFLite) per ottimizzarne l'esecuzione su hardware privo di GPU dedicata. Le differenze nei tempi di inferenza tra modelli della stessa famiglia riflettono le specifiche scelte architetturali: ad esempio, YOLOv9s presenta tempi superiori a YOLOv8s in ragione dell'introduzione dei meccanismi PGI (*Programmable Gradient Information*) e GELAN (*Generalized Efficient Layer Aggregation Network*), che aumentano la capacità rappresentativa del modello a fronte di un moderato incremento del costo computazionale. I tempi includono la fase di inferenza sul singolo frame; preprocessing e postprocessing sono gestiti dall'applicazione Python containerizzata tramite RabbitMQ RPC.

La sinergia tra l'architettura leggera del modello e la sua implementazione ottimizzata su hardware edge fa sì che YOLOv9s possa garantire sia un'inferenza ad alta velocità che prestazioni affidabili per le funzioni di sicurezza critiche in contesti produttivi. Questa specifica configurazione offre una soluzione concreta per l'identificazione automatizzata dei pericoli e la risposta immediata, aspetti vitali per minimizzare i rischi in ambienti come i forni elettrici ad arco, dove la rapidità nel prendere decisioni è essenziale. La possibilità di estendere il sistema a molteplici dispositivi assicura flessibilità nelle operazioni industriali, garantendo la sorveglianza continua dei principali parametri di sicurezza.

**Tab.2** -Tempi di inferenza ottenuti nei task di object detection / Inference times for the object detection tasks.

Detector Type	Model	Inference Time (ms)
Two-stage	Cascade R-CNN	16,56
	Faster R-CNN	25,5
	Mask R-CNN	12,4
Single-stage	YOLOv11s	5
	YOLOv11m	6,6
	YOLOv11x	12,8
	YOLOv10s	6,1
	YOLOv10m	7,2
	YOLOv10x	12
	YOLOv9s	9,8
	YOLOv9m	8,1
	YOLOv9e	15,5
	YOLOv8s	3,9
	YOLOv8m	6
	YOLOv8x	14,7
	RetinaNet	11,8
	RT-DETR-l	22,2
	RT-DETR-x	28,4

## RISULTATI DI RICONOSCIMENTO E DETEZIONE

Oltre alla velocità di inferenza, è stata valutata l'accuratezza dei modelli, verificando che le prestazioni di rilevamento fossero adeguate ai requisiti di sicurezza operativa in quasi real-time tipici dell'industria siderurgica. Nel complesso, i modelli analizzati hanno mostrato risultati molto soddisfacenti, come sintetizzato di seguito.

1. *Rilevamento delle Persone:* Per questa attività, finalizzata alla tutela dei lavoratori nelle aree ad alto rischio dell'EAF, YOLOv9s si è distinto come il modello più efficace. Esso ha raggiunto un mAP50-95 pari a 0.6174, con valori elevati di mAP50 (0.9610), precision (0.9375) e recall (0.8906), garantendo al contempo un tempo di inferenza di soli 9.8 ms per immagine. Que-

sto equilibrio tra accuratezza e rapidità lo rende particolarmente adatto ad applicazioni di sicurezza industriale. Altri modelli, come RT-DETR-l e RT-DETR-x, hanno mostrato buone prestazioni in termini di accuratezza, ma con tempi di risposta più elevati, meno compatibili con scenari operativi urgenti. I rilevatori a due stadi, caratterizzati da inferenza più lenta e minore precisione, risultano invece meno idonei per contesti industriali critici.

2. *Risultati del Rilevamento della posizione delle pinze degli elettrodi:* Nel monitoraggio delle pinze degli elettrodi, essenziale per il corretto e sicuro funzionamento dell'EAF, YOLOv9s ha nuovamente ottenuto le migliori prestazioni. Il modello ha registrato un

mAP50-95 di 0.8616, con valori prossimi all'unità per mAP50, mAP75, precision e recall, e un tempo di inferenza di 9.8 ms. Tali risultati evidenziano un'elevata affidabilità nel localizzare le pinze con un numero minimo di falsi positivi. Sebbene altri modelli abbiano mostrato prestazioni competitive, YOLOv9s ha garantito il miglior compromesso tra accuratezza e velocità, risultando la scelta più adatta per il monitoraggio in quasi real-time.

3. *Rilevamento del fumo*: Per il rilevamento del fumo è stato adottato un approccio di classificazione, utilizzando YOLO11s. Il modello ha raggiunto un'accu-

rezza del 92.55% e un punteggio di fitness (un indice globale in [0;1] utilizzato dai modelli YOLO per quantificare la bontà delle prestazioni di un modello) pari a 0.96 con precision di 0.98, recall di 0.97 e F1-score di 0.98, confermando un'elevata capacità di identificare correttamente la presenza di fumo con un numero contenuto di falsi negativi. Sebbene le prestazioni siano complessivamente soddisfacenti, rimane un margine di miglioramento, particolarmente rilevante in contesti industriali critici in cui l'individuazione precoce delle anomalie è fondamentale per la sicurezza.

**Tab.3** -Risultati ottenuti nel task di rilevamento delle persone / Results achieved on the person detection task.

Detector Type	Model	mAP50-95	mAP50	mAP75	Precision	Recall	F1 score
Two-stage	Cascade R-CNN	0,46	0,86	0,43	0,45	0,54	0,49
	Faster R-CNN	0,49	0,90	0,43	0,47	0,56	0,51
	Mask R-CNN	0,31	0,76	0,14	0,31	0,44	0,36
Single-stage	YOLOv11s	0,60	0,96	0,67	0,94	0,88	0,91
	YOLOv11m	0,59	0,96	0,63	0,94	0,88	0,91
	YOLOv11x	0,59	0,95	0,66	0,93	0,89	0,91
	YOLOv10s	0,58	0,95	0,62	0,94	0,88	0,91
	YOLOv10m	0,57	0,92	0,63	0,93	0,84	0,88
	YOLOv10x	0,60	0,95	0,66	0,96	0,85	0,90
	YOLOv9s	0,62	0,96	0,71	0,94	0,89	0,91
	YOLOv9m	0,60	0,96	0,66	0,94	0,89	0,91
	YOLOv9e	0,58	0,95	0,60	0,93	0,87	0,90
	YOLOv8s	0,58	0,95	0,67	0,90	0,88	0,89
	YOLOv8m	0,61	0,97	0,68	0,95	0,90	0,92
	YOLOv8x	0,60	0,95	0,68	0,94	0,88	0,91
	RetinaNet	0,51	0,93	0,50	0,49	0,58	0,53
	RT-DETR-l	0,60	0,95	0,71	0,94	0,91	0,92
	RT-DETR-x	0,60	0,96	0,68	0,92	0,92	0,92

**Tab.4** - Risultati ottenuti nel task di rilevamento delle pinze gli elettrodi / Results obtained on the electrodes clamp detection task.

Detector Type	Model	mAP50-95	mAP50	mAP75	Precision	Recall	F1 score
Two-stage	Cascade R-CNN	0,73	0,98	0,98	0,56	0,60	0,58
	Faster R-CNN	0,75	0,99	0,96	0,59	0,63	0,61
	Mask R-CNN	0,61	0,97	0,76	0,47	0,54	0,50
Single-stage	YOLOv11s	0,83	1,00	1,00	1,00	1,00	1,00
	YOLOv11m	0,83	0,99	0,99	1,00	0,99	0,99
	YOLOv11x	0,82	0,99	0,99	1,00	0,99	0,99
	YOLOv10s	0,86	1,00	1,00	1,00	0,99	1,00
	YOLOv10m	0,79	0,97	0,96	0,93	0,88	0,91
	YOLOv10x	0,85	0,99	0,99	0,99	0,98	0,99
	YOLOv9s	0,86	1,00	1,00	1,00	1,00	1,00
	YOLOv9m	0,84	0,99	0,99	1,00	0,98	0,99
	YOLOv9e	0,79	0,99	0,99	0,99	0,99	0,99
	YOLOv8s	0,85	1,00	0,99	1,00	1,00	1,00
	YOLOv8m	0,84	1,00	0,99	1,00	0,99	1,00
	YOLOv8x	0,86	0,99	0,99	1,00	0,99	0,99
	RetinaNet	0,74	0,99	0,95	0,59	0,64	0,61
	RT-DETR-l	0,83	0,99	0,99	1,00	0,99	0,99
	RT-DETR-x	0,84	0,99	0,99	0,98	0,99	0,98

## CONCLUSIONI

In questo lavoro è stata presentata EAFvision, una pipeline progettata per la valutazione e l'ottimizzazione di modelli di visione artificiale allo stato dell'arte applicati a compiti di rilevamento critici per la sicurezza nei forni elettrici ad arco. I risultati sperimentali hanno evidenziato come le architetture della famiglia YOLO, e in particolare YOLOv9s, si collochino tra le soluzioni più efficaci in termini di accuratezza, robustezza e velocità di inferenza, requisiti fondamentali per applicazioni industriali in tempo reale. Tuttavia, questa versione di EAFvision che sfrutta il framework Ultralytics non è stata ancora integrata nella linea di produzione e ulteriori valutazioni sul funzionamento

dovranno essere fatte in tale frangente.

YOLOv9s si è affermato come una delle soluzioni più efficaci per attività chiave quali il rilevamento del personale, il monitoraggio delle pinze degli elettrodi e l'identificazione del fumo, raggiungendo risultati comparabili o superiori agli altri modelli testati e garantendo un eccellente compromesso tra precisione e tempi di risposta. Il confronto con modelli concorrenti, tra cui YOLOv8, RT-DETR e i rilevatori a due stadi (Faster R-CNN, Mask R-CNN e Cascade R-CNN), ha confermato che le architetture YOLO risultano particolarmente adatte a contesti operativi critici, dove l'elaborazione rapida delle informazioni è determinante per prevenire situazioni di rischio. In particolare, le pre-

stazioni di YOLOv9s nel rilevamento del personale in aree pericolose evidenziano il suo potenziale nel supportare interventi tempestivi e nel ridurre il rischio di incidenti.

Il presente studio presenta alcune limitazioni che è opportuno riconoscere. Il dataset relativo al task di identificazione del fumo, composto da 429 immagini di training, è il più ridotto tra quelli considerati e potrebbe limitare la capacità di generalizzazione del modello a condizioni non rappresentate. Inoltre, tutti i dati sono stati acquisiti presso un singolo impianto EAF: la trasferibilità dei modelli ad altri contesti produttivi con caratteristiche impiantistiche o operative diverse rimane da verificare. Dal punto di vista della valutazione, i modelli sono stati testati su immagini statiche; la consistenza temporale delle predizioni su flussi video continui, rilevante per un sistema di monitoraggio in quasi real-time, non è stata oggetto di analisi sistematica.

Nel complesso, questo studio sperimentale dimostra il forte potenziale delle tecnologie di intelligenza artificiale e visione artificiale nel migliorare sia la sicurezza sia l'efficienza operativa nel settore siderurgico. Le prospettive future includono l'ulteriore ottimizzazione della pipeline

EAFvision, l'integrazione di nuove fonti sensoriali e l'estensione a ulteriori task critici per la sicurezza e il controllo di processo. Tali sviluppi mirano alla realizzazione di sistemi di monitoraggio intelligenti e pienamente automatizzati, capaci di supportare in modo affidabile la produzione dell'acciaio in ambienti industriali complessi e ad alto rischio.

## RINGRAZIAMENTI

Il lavoro descritto nel presente articolo è stato sviluppato nell'ambito del progetto dal titolo "*Remote expert virtual system enhancing human management capabilities that favors preservation, transfer, and continuous evolution of knowledge for steelmaking operations*" (Ref. ISteel-Expert, Grant Agreement No. 101112102), finanziato dall'Unione Europea tramite il "Research Fund for Coal and Steel (RFCS)", che gli Autori ringraziano. La responsabilità esclusiva delle questioni trattate nel presente lavoro è degli autori; l'Unione Europea non è responsabile per l'uso che può essere fatto delle informazioni ivi contenute.

## BIBLIOGRAFIA

- [1] H. NORDLOF, B. WIITAVAARA, U. WINBLAD, K. WIJK, R. WESTERLING Safety culture and reasons for risk-taking at a large steel-manufacturing company: Investigating the worker perspective, *Safety Science*, vol. 73, pp. 126–135, 2015.
- [2] R. LAN, I. AWOLUSI, J. CAI Computer vision for safety management in the steel industry, *AI*, vol. 5, no. 3, pp. 1192–1215, 2024
- [3] S. REN, K. HE, R. GIRSHICK, J. SUN Faster r-cnn: Towards real-time object detection with region proposal networks, *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1137–1149, 2016.
- [4] K. HE, G. GKIOXARI, P. DOLLAR, R. GIRSHICK Mask r-cnn, *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [5] Z. CAI, N. VASCONCELOS Cascade r-cnn: Delving into high quality object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 6154–6162.
- [6] G. JOCHER, A. CHAURASIA, J. QIU Ultralytics yolov8, 2023. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [7] C.-Y. WANG AND H.-Y. M. LIAO Yolov9: Learning what you want to learn using programmable gradient information, *European conference on computer vision*. Springer Nature Switzerland, 2024.
- [8] L. L. AO WANG, HUI CHEN Yolov10: Real-time end-to-end object detection, *arXiv preprint arXiv:2405.14458*, 2024
- [9] G. JOCHER, J. QIU Ultralytics yolo11, 2024. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [10] T.-Y. LIN, P. GOYAL, R. GIRSHICK, K. HE, P. DOLLAR Focal loss for dense object detection, *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.
- [11] N. CARION, F. MASSA, G. SYNNAEVE, N. USUNIER, A. KIRILLOV, S. ZAGORUYKO End-to-end object detection with transformers, *European conference on computer vision*. Springer, 2020, pp. 213–229.
- [12] Y. ZHAO, W. LV, S. XU, J. WEI, G. WANG, Q. DANG, Y. LIU, J. CHEN Detsr beat yolos on real-time object detection, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, pp. 16 965–16 974
- [13] Y. WU, A. KIRILLOV, F. MASSA, W.-Y. LO, R. GIRSHICK Detectron2, <https://github.com/facebookresearch/detectron2>, 2019.
- [14] I. LOSHCHILOV, F. HUTTER Decoupled weight decay regularization, in *International Conference on Learning Representations*, arXiv preprint arXiv:1711.05101, 2017.
- [15] R. PADILLA, S. NETTO, E. DA SILVA, A survey on performance metrics for object-detection algorithms. In: *2020 international conference on systems, signals and image processing (IWSSIP)*. IEEE, 2020. p. 237–242.

# Digital transformation in steel manufacturing: an AI-driven monitoring and safety framework for EAF operations

The iSteel-Expert project introduces an innovative AI-based monitoring and safety system specifically designed for Electric Arc Furnace operations in steel plants. The system integrates a network of sensors and cameras with a real-time data collection platform, utilizing advanced deep learning models to automate the precise detection and classification of critical events and safety conditions within the furnace environment. This innovative approach significantly enhances both operational safety standards and overall process efficiency.

The monitoring framework focuses on several key operational aspects: person detection in the furnace area to assess human exposure and operator compliance with established procedures, slag door status monitoring to verify process compliance during the heat, continuous tracking of electrode arm movement through clamp position detection to identify anomalies that may require corrections or maintenance interventions, and detection of vapor, fumes, and flames to ensure environmental and operational safety.

A comprehensive dataset of approximately 5000 images was collected for model training and validation. The system demonstrated excellent performance, achieving accuracy exceeding 90% for furnace tilt and smoke detection, with high precision, recall, and F1-scores in object detection tasks.

**KEYWORDS:** EAF; SAFETY IMAGE PROCESSING; DEEP LEARNING; ARTIFICIAL INTELLIGENCE.

[TORNA ALL'INDICE >](#)