



BANCA D'ITALIA
EUROSISTEMA

Mercati, infrastrutture, sistemi di pagamento

(Markets, Infrastructures, Payment Systems)

Reti neurali siamesi per la rilevazione dei difetti di stampa
delle banconote

di Katia Boria, Andrea Luciani, Sabina Marchetti e Marco Viticoli



BANCA D'ITALIA
EUROSISTEMA

Mercati, infrastrutture, sistemi di pagamento

(Markets, Infrastructures, Payment Systems)

Approfondimenti
(Research Papers)

Reti neurali siamesi per la rilevazione dei difetti di stampa
delle banconote

di Katia Boria, Andrea Luciani, Sabina Marchetti e Marco Viticoli

Numero 34 – maggio 2023

I lavori pubblicati nella collana “Mercati, infrastrutture, sistemi di pagamento” presentano documentazioni e studi su aspetti rilevanti per i compiti istituzionali della Banca d’Italia in tema di monitoraggio dei mercati finanziari e del sistema dei pagamenti, nonché di sviluppo e gestione delle relative infrastrutture. L’intento è quello di contribuire alla diffusione della conoscenza su questi argomenti e di favorire il dibattito tra le istituzioni, gli operatori economici, i cittadini.

I lavori pubblicati riflettono le opinioni degli autori, senza impegnare la responsabilità dell’Istituto.

La serie è disponibile online sul sito www.bancaditalia.it.

Copie a stampa possono essere richieste alla casella della Biblioteca Paolo Baffi: richieste.pubblicazioni@bancaditalia.it.

Comitato di redazione: STEFANO SIVIERO, LIVIO TORNETTA, GIUSEPPE ZINGRILLO, GUERINO ARDIZZI, PAOLO LIBRI, GIUSEPPE MARESCA, ONOFRIO PANZARINO, TIZIANA PIETRAFORTE, ANTONIO SPARACINO.

Segreteria: ALESSANDRA ROLLO.

ISSN 2724-6418 (online)
ISSN 2724-640X (stampa)

Banca d’Italia
Via Nazionale, 91 - 00184 Roma - Italia
+39 06 47921

Grafica e stampa a cura della Divisione Editoria e stampa della Banca d’Italia

RETI NEURALI SIAMESI PER LA RILEVAZIONE DEI DIFETTI DI STAMPA DELLE BANCONOTE

di Katia Boria*, Andrea Luciani**, Sabina Marchetti** e Marco Viticoli***

Sintesi

La produzione di banconote è un processo complesso, costituito da differenti fasi di stampa, nell'ambito delle quali si possono generare dei difetti che, se non adeguatamente monitorati, danno origine a scarti di produzione con impatti significativi sulla produttività e sui costi. Il presente lavoro propone un approccio innovativo per identificare i difetti in fase di produzione di banconote, avvalendosi di metodi di apprendimento "one-shot". Questi metodi permettono di utilizzare un numero limitato di osservazioni al fine di addestrare una rete neurale siamese per renderla in grado di individuare le caratteristiche di similarità tra coppie di esempi. La rete viene quindi impiegata per identificare i difetti su nuove immagini di banconote, confrontandole con esempi di riferimento. L'approccio individuato consente la corretta identificazione dei difetti oggetto dello studio, anche a fronte di una limitata quantità di osservazioni, gettando le basi per lo sviluppo di una soluzione per il riconoscimento e la classificazione "intelligente" dei difetti delle banconote.

Classificazione JEL: C45, L15, L69, O39.

Parole chiave: produzione di banconote, intelligenza artificiale, reti neurali, apprendimento one-shot, qualità.

Abstract

The production of banknotes is a complex process, composed of different printing steps, in which various kinds of defects can be generated which, if not adequately monitored, can lead to production waste, significantly impacting productivity and costs. This paper proposes a novel approach for identifying defects during banknote production using 'one-shot learning' methods. These methods rely on a small number of observations to train a Siamese neural network to reproduce the similarities between pairs of samples. The network can then identify defects in new banknote images by comparing them to benchmark samples. The proposed approach allows the correct identification of some specific defects on banknotes, even with limited training data, laying the foundation for the development of a solution for the recognition and intelligent classification of defects on banknotes.

* Banca d'Italia, Dipartimento Informatica.

** Banca d'Italia, Dipartimento Economia e statistica.

*** Banca d'Italia, Dipartimento Circolazione e pagamenti al dettaglio.

INDICE

1. Introduzione	7
2. I difetti nella stampa di banconote e la variabilità di produzione	7
3. Metodologia	9
3.1 Dati	9
3.2 Deep learning e reti neurali convoluzionali	9
3.3 Reti neurali convoluzionali per lo studio delle banconote	12
3.4 One-shot learning e reti neurali siamesi	13
4. Reti siamesi per la rilevazione dei difetti di stampa di tipo bite	15
4.1 Architettura e addestramento dei modelli	15
4.2 Risultati	18
5. Conclusioni e possibili estensioni	21
Riferimenti bibliografici	23
Appendice: Estensione a difetti bite-L	28

1. Introduzione

In analogia ad altri settori industriali, la produzione delle banconote si avvale di un articolato sistema di controlli di qualità. Nell'Eurosistema, tali controlli sono condotti durante l'intero processo produttivo (materie prime, semilavorati e banconote finite) nel rispetto delle norme ISO di riferimento e mediante rigorose procedure definite dalla Banca Centrale Europea (BCE) e armonizzate tra le varie stamperie per garantire un elevato grado di omogeneità dei biglietti in euro. Con particolare riferimento alle verifiche di qualità sulle banconote finite, la numerosità e la tipologia dei difetti di stampa rappresenta un parametro discriminante per la conformità di ciascun lotto di produzione. Oggi, la valutazione delle difettosità è effettuata dagli operatori attraverso l'esame visivo dei biglietti. Pertanto, sebbene esista una articolata serie di presidi finalizzati a garantire l'applicazione di criteri oggettivi, l'esperienza e la sensibilità degli addetti giocano un ruolo determinante nel processo valutativo, che potrebbe quindi risentire di fattori di soggettività.

In linea di principio, l'automatizzazione dei controlli di qualità dei prodotti industriali contribuisce a rendere i processi più precisi ed efficienti. L'intelligenza artificiale, e in particolare le reti neurali nell'ambito del *deep learning*, sono sfruttate con profitto in tal senso, grazie alla loro abilità di rappresentare e descrivere sistemi caratterizzati da alta complessità e variabilità. In fase di stampa, la valutazione e identificazione dei casi critici da parte delle reti neurali consente di ridurre i costi operativi, demandando agli specialisti dei controlli di qualità l'analisi di dettaglio dei potenziali difetti riscontrati e delle loro cause.

Il presente lavoro riporta quanto osservato nell'ambito di uno studio esplorativo sull'uso di reti neurali per il riconoscimento e la classificazione "intelligente" di una particolare tipologia di difetti delle banconote nell'ambito dei controlli di qualità. La Sezione 2 fornisce una panoramica essenziale del processo di controllo della qualità per la rilevazione di eventuali difetti di stampa sulle banconote, a valle della fase di produzione. Nella Sezione 3 si delineano quindi le principali caratteristiche tecniche delle reti neurali *siamesi* - ovvero la classe di modelli di *deep learning* utilizzata nel presente studio - nonché dell'insieme di immagini utilizzate per l'addestramento e la calibrazione. I principali aspetti relativi all'addestramento dei modelli e i risultati conseguiti utilizzando le reti siamesi sono riportati e discussi nella Sezione 4. La Sezione 5, infine, traccia le possibili direzioni di sviluppo e approfondimento, di carattere sia metodologico che operativo.

2. I difetti nella stampa di banconote e la variabilità di produzione

Nell'ambito del processo produttivo rigorosi controlli di qualità sono effettuati sulle banconote prodotte e sui semilavorati, nel rispetto delle decisioni di riferimento in materia di qualità della Banca Centrale Europea (BCE) e delle prescrizioni del Sistema di Gestione Integrato Qualità, Ambiente e Salute e sicurezza sul lavoro.

L'insieme di questi controlli fornisce importanti informazioni sulle potenziali criticità del processo e sulle difettosità del prodotto finale che possono essere proficuamente utilizzate per intervenire sul processo stesso al fine di ridurre gli scarti di produzione già nelle prime fasi di stampa e le difettosità anche marginali del prodotto finito.

La verifica di numerosi parametri è stata già automatizzata tramite l'introduzione di sistemi ottici di misura. Tuttavia, i controlli visivi di accettazione sulle banconote finite sono ancora affidati a personale altamente qualificato, che li effettua manualmente su un campione rappresentativo di ogni lotto di produzione. I difetti sono quindi individuati e classificati sulla base dei documenti della BCE e di campioni di riferimento. La classificazione avviene per tipo e dimensione del difetto, nonché per zona della superficie in cui i diversi difetti si possono trovare. La valutazione finale di conformità del lotto di produzione dipende quindi dal numero e dalla classificazione dei difetti trovati sul campione rappresentativo del lotto. Tale processo si basa su una serie di regole e sul confronto con campioni di riferimento, in maniera analoga a ciò che viene eseguito in modo automatico dalle tecniche di *deep learning*.

Per la sperimentazione descritta nel presente lavoro, si è scelto di realizzare una applicazione pilota per l'individuazione dei difetti sulle banconote del 50€ della serie Europa¹ attualmente in produzione presso il Servizio Banconote della Banca d'Italia. In questa prima fase, lo studio si è concentrato su una specifica tipologia di difetto presente sulla bandiera europea, il cosiddetto *bite*, che consiste nella mancanza di inchiostro su un fondo omogeneo, e può avere forma, dimensioni e posizione variabile. La scelta è stata determinata dalla rilevanza e numerosità di tale tipologia di difetto nei lotti di produzione esaminati al momento della sperimentazione.

La variabilità delle banconote è stato un altro importante fattore di analisi di cui si è dovuto tener conto in fase sperimentale. Le banconote infatti sono il risultato di diverse fasi di stampa, quindi possono presentare tra loro differenze apprezzabili ma che, rientrando nella normale variabilità produttiva, non devono essere considerate dei difetti. In particolare:

- 1) la posizione degli elementi all'interno della banconota può variare pur mantenendosi entro fissati margini di tolleranza;
- 2) la densità e il colore degli inchiostri può variare nell'ambito dei lotti di produzione determinando tonalità e intensità differenti;
- 3) la fase di acquisizione sia manuale che soprattutto automatica può introdurre ulteriore variabilità delle immagini.

Oltre alla variabilità di prodotto, è stato necessario analizzare anche la variabilità dei difetti da identificare e classificare, diversi per dimensioni, posizione e intensità di colore.

Alla luce di tali evidenze, si è inizialmente proceduto con l'acquisizione manuale delle immagini ad alta risoluzione (~2656x1467 pixel, 24 bit colore, modello RGB) della banconota, cercando di minimizzare i difetti di acquisizione. Successivamente, al fine di incrementare in maniera significativa il numero di campioni da processare, si sono raccolte le immagini acquisite in modo automatico durante il processo produttivo, che presentavano una risoluzione più bassa (~675x388 pixel, 24 bit colore, modello RGB). In entrambi i casi l'acquisizione delle immagini delle banconote ha richiesto l'intervento dell'operatore, anche se nel secondo caso la procedura è stata decisamente più rapida. L'acquisizione manuale ha consentito d'altra parte una più attenta selezione dei difetti da acquisire, rappresentativi di un arco temporale di produzione più ampio (alcune settimane) e di una maggiore variabilità produttiva. Allo stato attuale, i due diversi approcci di acquisizione sono da

¹ La serie Europa è la seconda serie delle banconote Euro, introdotte dal 2013 al 2019.

tenere in considerazione per i loro effetti sulla disponibilità, entità e variabilità del campione di immagini, in fase di apprendimento del modello utilizzato.

3. Metodologia

L'applicazione proposta si focalizza sulla rilevazione dei difetti di stampa di tipo *bite* mediante l'uso di reti neurali siamesi. La presente sezione introduce il lettore agli aspetti tecnici dell'analisi condotta. Si descrivono brevemente: i) la tipologia di rappresentazione dei dati in formato immagine, secondo serie ordinate di matrici dette *tensori*; ii) le reti neurali convoluzionali utilizzate per l'elaborazione di tensori, i limiti di tali modelli, di carattere generale e specifici all'applicazione su immagini di banconote; iii) la specializzazione del *deep learning*, nota come *one-shot learning*, e la relativa classe delle reti neurali siamesi.

3.1 Dati

Da un punto di vista tecnico, le immagini appartengono alla tipologia dei dati in formato “non strutturato”. Ai fini dell'analisi statistica, ciascuna osservazione viene rappresentata ricorrendo ad una serie ordinata di tre matrici, o tensore di rango tre, che ne sintetizza le caratteristiche di interesse. In particolare, gli elementi di queste matrici corrispondono ai componenti grafici elementari dell'immagine rappresentata, ovvero i c.d. *pixel*². All'interno di una griglia così definita, ciascun elemento della matrice contiene il valore numerico del *pixel* corrispondente; questo valore, nel caso di colore a 24 bit, varia tra 0 e 255 e indica l'intensità del colore associato alla matrice secondo il suo posizionamento nella sequenza. Nel caso del modello RGB, la prima matrice rappresenta il c.d. *canale* del colore rosso (*Red*), la seconda matrice rappresenta il canale del colore verde (*Green*), la terza del colore blu (*Blue*)³. Ad esempio, secondo tale schema di rappresentazione un *pixel* blu si ottiene impostando a zero i relativi valori matriciali nel primo canale (*Red*) e nel secondo canale (*Green*), e impostando a 255 il valore matriciale nel terzo canale (*Blue*). Nel caso generale quindi, data una risoluzione di immagine di $N \times M$ *pixel* codificata secondo il modello RGB, l'immagine sarà numericamente rappresentata da un tensore di rango tre contenente $N \times M \times 3$ numeri interi, ognuno dei quali può come detto variare tra 0 e 255.

3.2 Deep learning e reti neurali convoluzionali

L'analisi avanzata di immagini si avvale in genere di tecniche di intelligenza artificiale e in particolare del *deep learning*, soprattutto nei problemi di riconoscimento o di classificazione automatica dell'oggetto rappresentato in immagine. Il *deep learning*, che è identificabile come una specializzazione del *machine learning*, estende il processo di apprendimento di una regola di rappresentazione dell'input a sequenze di strati di rappresentazione (Goodfellow, Bengio & Courville, 2016). L'approccio si fonda sul paradigma *data-driven*, secondo il quale l'informazione

² I *pixel* sono unità minime della superficie di un'immagine digitale.

³ Un altro modello frequentemente utilizzato è il cosiddetto HSV, per il quale i valori delle celle variano tra 0 e 360. Il primo canale corrisponde alla tinta (*Hue*). Il secondo canale identifica il grado di saturazione (*Saturation*). Il terzo canale fa riferimento alla luminosità del colore (*Color Value*).

necessaria alla comprensione di un dato fenomeno, e al corretto funzionamento del relativo modello⁴ di rappresentazione, è appresa a partire da un insieme di dati empirici già disponibili all'analista, detto insieme di *training*. Il cosiddetto addestramento del modello non richiede quindi la specificazione e la formalizzazione di un elevato numero di assunzioni sul fenomeno oggetto di analisi o sul processo generativo dei dati.

Le reti neurali sono classi di modelli per il *deep learning* in cui ogni componente interno di trasformazione dell'informazione (*layer*) identifica uno strato di rappresentazione dell'input. All'interno degli strati, l'elaborazione dell'informazione sfrutta le unità di base, dette neuroni. Nella loro formalizzazione più semplice, detta "completamente connessa", i neuroni ricevono, elaborano e trasmettono l'informazione che è veicolata dal *layer* iniziale (Input) a quello finale (Output), passando per strati intermedi nascosti, o *hidden layers*.

In una rete neurale con L strati nascosti, sia $\mathbf{y}^{l+1} \in \mathbb{R}^{n_{l+1}}$ la proiezione dell'osservazione in input $\mathbf{y}^0 \in \mathbb{R}^{n_0}$ prodotta dall' l -esimo (ellesimo) strato completamente connesso, $l=0, \dots, L$. Il vettore \mathbf{y}^{l+1} è ricavato attraverso la formula ricorsiva:

$$\mathbf{y}^{l+1} = g^l(\mathbf{w}^l \mathbf{y}^l + \mathbf{b}^l)$$

dove g^l è una generica *funzione di attivazione* che si applica alla trasformazione lineare del vettore in input $\mathbf{y}^l \in \mathbb{R}^{n_l}$. Il valore dei parametri, i.e. la matrice dei pesi $\mathbf{w}^l \in \mathbb{R}^{n_{l+1} \times n_l}$ e il vettore dei termini di distorsione $\mathbf{b}^l \in \mathbb{R}^{n_{l+1}}$, sono opportunamente calibrati durante il processo di apprendimento data un'inizializzazione casuale, e come detto, utilizzando un insieme di dati empirici già disponibili all'analista, detto insieme di *training*.

L'addestramento di un modello è quindi associato alla fase di "calibrazione", utile a validarne la *performance* su un nuovo insieme di osservazioni e a stabilire la corretta configurazione degli iper-parametri del modello, ovvero l'insieme delle sue caratteristiche generali fissate a priori dall'analista. La calibrazione di una rete neurale ne definisce quindi l'architettura, che include: i) la tipologia degli strati che regolano la trasmissione dell'informazione; ii) la profondità della rete neurale, i.e. il numero di strati che la caratterizzano; iii) il valore degli iper-parametri che ne regolano il funzionamento, come ad esempio il numero di neuroni o la funzione di attivazione utilizzati; iv) il processo di apprendimento impiegato per la calibrazione dei diversi elementi.

Le applicazioni di *image processing* ricorrono tipicamente a reti neurali contenenti strati di tipo convoluzionale (*Convolutional Neural Networks*, CNN)⁵. Le CNN si ispirano al meccanismo biologico di percezione visiva. Esse operano una riduzione della complessità del dato in input attraverso l'estrazione mirata di una sintesi che rende efficiente il processo di elaborazione. Tecnicamente gli strati convoluzionali elaborano l'informazione rappresentata da un tensore secondo un procedimento basato su filtri ricorsivi che scorrono lungo le dimensioni della superficie, vale a dire lungo gruppi di pixel adiacenti tra loro, o regione. A partire da ciascuna regione si estraggono delle proiezioni di sintesi dette *filter map*.

⁴ Il corretto funzionamento del modello è espresso dal valore di una funzione obiettivo, che ne guida il processo di addestramento.

⁵ LeCun *et al.* (1989).

Al fine di garantire la stabilità del sistema di elaborazione dell'informazione, gli strati convoluzionali si alternano a strati di raccolta (*pooling*) che sintetizzano e aggregano il contributo delle regioni all'interno di un *filter map* (Boureau *et al.*, 2010). Tipicamente, da ciascun elemento prodotto dall'operazione di convoluzione, gli strati di raccolta recepiscono o il valore medio ottenuto (*average pooling*) o il valore massimo (*max pooling*). L'operazione di raccolta genera quindi un tensore di dimensionalità ridotta i cui elementi costituiscono l'input per lo strato successivo.

La natura ricorsiva e a regioni dello schema di rilevamento implica che l'informazione relativa allo stesso pixel sia presente in più proiezioni. Ciò rende ciascun *filter map* molto sensibile a piccoli spostamenti di un elemento all'interno dell'area dell'immagine, e quindi dei pixel che concorrono a definirlo. Tale aspetto risulta essere particolarmente rilevante nel caso d'uso considerato, in quanto le banconote sono caratterizzate da un'organizzazione stabile degli elementi rappresentati sulla propria superficie, pur con una tollerata variabilità data dal loro posizionamento reciproco e dall'intensità del colore. L'uso di CNN per la classificazione nell'ambito dell'elaborazione delle immagini si distingue in approcci basati su segmentazione dell'input e riconoscimento dei diversi elementi (*object recognition*) e su assegnazione di un'unica etichetta all'osservazione nella sua interezza (*image recognition*). Il riconoscimento degli elementi può operare la segmentazione delle immagini in aree specifiche della superficie individuate come potenzialmente rilevanti, da trasmettere a modelli ad hoc per l'attribuzione di un'etichetta ai singoli elementi, o incorporare le due fasi all'interno di un'unica rete neurale (Redmon *et al.*, 2016). Le CNN per il riconoscimento dei diversi elementi di una immagine sono solitamente caratterizzate da architetture complesse, alle quali si associano costi computazionali rilevanti per le fasi di addestramento e di applicazione. La letteratura relativa alla classificazione delle immagini ha prodotto nel tempo CNN caratterizzate da elevate prestazioni su grandi volumi di dati, anche grazie alla raffinata complessità delle architetture (Krizhevsky *et al.*, 2017). In particolare, l'accuratezza del processo di classificazione è resa possibile da reti neurali molto profonde organizzate secondo strutture gerarchiche (Simonyan & Zisserman, 2014; Szegedy *et al.*, 2015; Yan *et al.*, 2015) e dalla parametrizzazione granulare degli strati, in cui la dimensione dei filtri ricorsivi approssima il singolo pixel (Zeiler & Fergus, 2014).

L'alta dimensionalità dell'input e il carattere ricorsivo dell'elaborazione dell'informazione fanno sì che le CNN impiegate per l'*image recognition* siano caratterizzate da un significativo numero di parametri da stimare, ai quali è attribuito un valore casuale in fase di inizializzazione. Il processo di addestramento di una CNN consiste nell'ottimizzazione globale del modello, attraverso l'adattamento progressivo di tali parametri. L'elevato numero di essi, associato ad architetture profonde e ramificate, fa sì che l'addestramento di una CNN richieda un'alta numerosità dell'insieme di *training*, così che il modello nel suo complesso apprenda la logica ottimale da applicare al processo di classificazione. Tale numerosità è di particolare rilievo per l'applicazione in oggetto.

La flessibilità di modelli caratterizzati da architetture complesse espone, come noto, gli esercizi di classificazione al cosiddetto *overfitting*, vale a dire la modellizzazione dell'errore irriducibile nei dati che penalizza la capacità di generalizzazione del modello e l'interpretabilità dei risultati. Per mitigare questo rischio e favorire la convergenza del processo di addestramento, alleggerendo al contempo la complessità computazionale dell'elaborazione, è possibile adottare una serie di accorgimenti. La presente applicazione, in particolare, ricorre a: i) metodi ad hoc per l'inizializzazione dei parametri in grado di attribuire valori iniziali all'interno di range ottimali (Domínguez, 2020), come ad esempio il metodo cosiddetto *Xavier* per la definizione dei parametri di una distribuzione uniforme (Glorot &

Bengio, 2010); ii) la calibrazione ottimale della routine di ottimizzazione, in genere appartenente alla categoria dello *Stochastic Gradient Descent* (SGD) per le CNN (Qian, 1999; Wijnhoven & De With, 2010); iii) la normalizzazione dell'input (*batch normalization*) per smussare i valori della funzione obiettivo e della derivazione del gradiente associato (Ioffe & Szegedy, 2015); iv) l'uso di strumenti di penalizzazione della complessità quali il *dropout* (Hinton *et al.*, 2012), vale a dire la disattivazione casuale di alcune connessioni tra i neuroni in fase di addestramento, e la regolarizzazione quadratica dei pesi negli strati convoluzionali (Yu *et al.*, 2008); v) la re-inizializzazione del valore dei parametri in caso di raggiungimento di un ottimo locale, per favorire l'esplorazione ottimale dello spazio delle soluzioni (Treadgold & Gedeon, 1996; Guo & Li, 2006).

3.3 Reti neurali convoluzionali per lo studio delle banconote

L'uso di CNN per l'analisi delle immagini di banconote apporta un notevole valore aggiunto principalmente dovuto a due fattori: la flessibilità dei modelli e la possibilità di elaborare i dati in formato immagine senza dover ricorrere all'uso di sensoristica per l'estrazione di variabili di sintesi, o *features* (Lee *et al.*, 2017). In letteratura e relativamente alle banconote, le applicazioni ricorrono in genere a modelli statistico-matematici e a tecniche di *machine learning* per il riconoscimento della denominazione della banconota (Grijalva *et al.*, 2010; Sharma *et al.*, 2012; García-Lamont *et al.*, 2013), del numero di serie (Feng *et al.*, 2014; Liu *et al.*, 2010; Wenhong *et al.*, 2010; Hasanuzzaman *et al.*, 2011), della valuta (Manikandan & Sumithra, 2015), dello stato di logoramento (Sun & Li, 2008; Daraee & Mozaffari, 2010; Mousavi *et al.*, 2015) o per la rilevazione di esemplari contraffatti (Darade *et al.*, 2016; Suresh *et al.*, 2016). Tecniche analitiche avanzate sono inoltre impiegate per il conteggio, il supporto operativo a soggetti diversamente abili e il monitoraggio della qualità del contante circolante da parte degli esercenti e delle stesse banche.

In letteratura recente si evidenziano anche pubblicazioni in cui sono utilizzate le CNN per il riconoscimento delle banconote (Jadhav *et al.*, 2019; Zhang *et al.*, 2019; Jang *et al.*, 2020; Park *et al.*, 2020; Veeramsetty *et al.*, 2020) e per la rilevazione di esemplari contraffatti (Kamble *et al.*, 2019; Sawant *et al.*, 2022), ottenendo un aumento generale della qualità della performance. Per superare le criticità associate all'uso di CNN, come ad esempio l'elevato volume di osservazioni necessarie per la convergenza del modello verso l'ottimo globale, e la tendenza delle stesse all'*overfitting* data dalle caratteristiche di variabilità associate alle banconote, le CNN sono state in alcuni casi sviluppate secondo approcci di *adversarial learning* per la generazione di immagini sintetiche (Ali *et al.*, 2019; Desai *et al.*, 2021; Khemiri *et al.*, 2022) o di *transfer learning*, vale a dire approcci che sfruttano modelli già addestrati a svolgere operazioni generali, per poi operarne la specializzazione grazie ad un processo di *training* in genere meno intensivo (Laavanya & Vijayaraghavan, 2019; Linkon *et al.*, 2020; Aseffa *et al.*, 2022).

Sono rintracciabili inoltre applicazioni riferite alla fase di produzione (Ke *et al.*, 2016; Pham *et al.*, 2017). In particolare, l'applicazione condotta da Ke e co-autori si concentra sulla rilevazione di difetti analoghi al *bite* – la stessa classe di difetti utilizzata in questo lavoro – sviluppando una CNN a partire da un insieme di training le cui caratteristiche principali non sono però pienamente esposte dagli autori.

L'analisi esplorativa delle metodologie proposte dalla letteratura per l'elaborazione di immagini attraverso tecniche di *deep learning* ha evidenziato come, ai fini dell'applicazione in oggetto, l'addestramento di una CNN sufficientemente profonda presenterebbe una serie di limiti, dati dalla già citata variabilità nella disposizione degli elementi grafici, dei difetti sulla superficie dell'immagine, e dalla limitata disponibilità di osservazioni per l'addestramento. Per mitigare la relativamente scarsa disponibilità di osservazioni, fissata un'architettura, si è considerata l'estensione dell'insieme di *training* ricorrendo a soluzioni di *data augmentation* (Wang & Perez, 2017). Tra le estensioni considerate in fase preliminare, si citano l'alterazione delle immagini esistenti attraverso tecniche di rotazione, ricampionamento e trasposizione, e la generazione di osservazioni sintetiche ottenute perturbando le immagini disponibili o addestrando un modello di supporto per la generazione di immagini (Goodfellow *et al.*, 2020).

3.4 One-shot learning e reti neurali siamesi

L'analisi esplorativa condotta ha evidenziato come la *data augmentation*, nel caso di specie considerato, vada valutata alla luce del *trade-off* costituito dall'instabilità del processo di addestramento in presenza di volumi limitati di osservazioni, e le possibili forme di distorsione associate all'aumento sintetico dei dati. In particolare, la *data augmentation* tende ad esporre la rete neurale a due tipi di criticità. Da un lato, immagini addizionali che sovra-rappresentino l'occorrenza di difetti di stampa con determinate caratteristiche rendono il modello esposto all'*overfitting* sulla tipologia del difetto: potrebbe ad esempio non essere in grado di rilevare correttamente la presenza di una macchia di inchiostro con caratteristiche diverse da quelle osservate in fase di training. Dall'altro, la generazione di nuovi difetti, non osservati, potrebbe introdurre forme di distorsione algoritmica rilevanti, nel caso in cui questi fossero inverosimili ma determinanti nella definizione della logica appresa dal modello.

Per arginare i limiti dati dalla scarsa rappresentatività dei difetti disponibili e dal combinato delle caratteristiche di variabilità di questi e delle immagini di banconote, l'esercizio di addestramento e calibrazione del modello è stato condotto facendo ricorso ad architetture CNN sviluppate secondo un approccio di tipo *one-shot learning* (Fei *et al.*, 2006).

L'apprendimento *one-shot* costituisce una specializzazione dell'apprendimento statistico, avente l'obiettivo di emulare la capacità della mente umana di associare tra loro entità sconosciute al soggetto, in base a criteri di similarità. L'approccio si propone quindi di riprodurre l'abilità degli esseri viventi più evoluti di operare il confronto visivo a partire dalla conoscenza di una verità di base conferita dall'esperienza. Data l'immagine di una generica entità – ad esempio un oggetto, un volto umano o un panorama – e una gamma di possibili categorie di appartenenza, l'assegnazione dell'immagine a una di queste categorie avviene a partire dalla simultanea identificazione dei punti in comune e di quelli di differenza con immagini rappresentative di ciascuna classe.

Nella sua formulazione più rigorosa l'apprendimento *one-shot* prevede che l'insieme di *training* sia costituito da un numero di osservazioni prototipali pari al numero di categorie oggetto di interesse. Tuttavia, in letteratura ci si riferisce in genere all'apprendimento *one-shot* nella sua accezione

maggiormente estesa⁶ denominata *k-shot*, in cui si considera un certo numero di esempi per classe, differenziandolo poi dall'apprendimento *zero-shot* (Palatucci *et al.*, 2009) in cui non si dispone di alcuna immagine di esempio per classe in fase di addestramento⁷.

Per il *one-shot learning* ci si può avvalere della classe di reti neurali di tipo *siamese* (Koch *et al.*, 2015). Le reti neurali siamesi sono caratterizzate dalla presenza di due sotto-architetture parallele (i.e. rami) identiche tra loro per struttura e valori dei parametri (come ad esempio in verde in figura 1). Ciascun ramo di una rete siamese elabora il rispettivo elemento della coppia di osservazioni che ne costituisce l'input: un'osservazione di “verifica”, per la quale si intende stabilire l'appartenenza a una data categoria, ed una di “supporto”, che rappresenta una data classe. L'elaborazione proietta in parallelo la coppia di osservazioni fino a produrne un “segnale”, vale a dire una rappresentazione dell'input sintetica e separabile. L'intuizione è la seguente: due osservazioni appartenenti alla stessa categoria sono proiettate dalle architetture parallele su segnali simili tra loro; viceversa, osservazioni appartenenti a categorie differenti, saranno caratterizzate da segnali dissimili. I segnali di verifica e di supporto, rispettivamente $\mathbf{y}^{Sig,v}, \mathbf{y}^{Sig,s} \in \mathbb{R}^m$, sono quindi combinati ed elaborati da uno strato “di confronto” (figura 1, in arancione), che ne quantifica la similarità mediante la rappresentazione $\mathbf{y}^{L'} \in \mathbb{R}^{n_{L'}}$, ottenuta dalla trasformazione:

$$\mathbf{y}^{L'} = g^{Sim}(\mathbf{w}^{Sim} Sim(\mathbf{y}^{Sig,v}, \mathbf{y}^{Sig,s}) + \mathbf{b}^{Sim})$$

dove g^{Sim} e Sim sono rispettivamente una funzione di attivazione e una funzione di similarità, $\mathbf{w}^{Sim} \in \mathbb{R}^{n_{L'} \times m}$ è la matrice dei pesi e $\mathbf{b}^{Sim} \in \mathbb{R}^m$ è il vettore dei termini di distorsione, sempre in riferimento allo strato di confronto. Nel caso generale $L' \leq (L + 1)$, ovvero dallo strato di confronto potrebbe innestarsi una ulteriore sequenza di strati di rappresentazione. Nel caso più semplice, considerato dalla presente applicazione, $L' = L + 1$, ovvero la rappresentazione ottenuta in uscita dallo strato di confronto è definitivamente l'output della rete siamese, tale per cui, nel nostro caso di studio binario, $n_{L'} = 1$ e g^{Sim} è la funzione sigmoide che proietta $y^{L+1} = \mathbf{y}^{L'}$ nell'intervallo $[0,1]$. In fase di addestramento, siano N le coppie di immagini dell'insieme di *training*; ciascuna coppia è etichettata come contenente immagini – una di verifica e una supporto – appartenenti alla stessa classe ($c_i = 1$) o a classi distinte ($c_i = 0$). I valori $y_1^{L+1}, \dots, y_N^{L+1}$ prodotti dalla rete siamese corrispondono pertanto alle probabilità che le N immagini di verifica appartengano alla stessa classe delle rispettive immagini di supporto. L'addestramento del modello consiste quindi nella minimizzazione della seguente *funzione di perdita*:

$$FP(\mathbf{y}^{L+1}) = -\frac{1}{N} \sum_{i=1, \dots, N} c_i \log(y_i^{L+1}) + (1 - c_i) \log(1 - y_i^{L+1})$$

dove c_i assume come detto valore 1 quando le osservazioni di verifica e di supporto appartengono alla medesima classe, 0 altrimenti, e $y_i^{L+1} = Prob(c_i = 1), i = 1, \dots, N$.

⁶ L'apprendimento *one-shot* va distinto dal “trasferimento *one-shot*”, da intendersi come specializzazione dell'apprendimento per trasferimento (*transfer learning*).

⁷ Sono presenti in letteratura numerosi esempi di applicazione di reti neurali basate su apprendimento *zero-shot*, nonché nell'ambito dell'elaborazione del linguaggio naturale. Tra le applicazioni più frequenti vi sono gli esercizi di classificazione in condizioni di estremo sbilanciamento delle classi (Ochal *et al.*, 2021), nonché le attività di riconoscimento facciale, alle quali è richiesto di garantire un certo livello di *performance* in condizioni di elevata variabilità, ad esempio nell'espressione facciale, nella luminosità o nello sfondo.

4. Reti siamesi per la rilevazione dei difetti di stampa di tipo *bite*

4.1 Architettura e addestramento dei modelli

Nel caso specifico di sviluppo di una rete siamese per la rilevazione di difetti *bite* sull'elemento grafico raffigurante la bandiera europea, si fa riferimento a un'architettura essenziale costituita da una coppia di reti gemelle così come rappresentata in figura 1. Ciascuna rete è caratterizzata da tre strati convoluzionali (*Conv2D*) cui seguono tre strati di raccolta (*MaxPooling2D*), sui quali si innesta una sequenza di strati completamente connessi (*Dense*) per la derivazione del vettore di segnale. Questi ultimi sono combinati secondo la metrica di distanza assoluta e utilizzati per derivare il valore di output y^{L+1} , corrispondente alla probabilità che la coppia appartenga alla stessa categoria⁸.

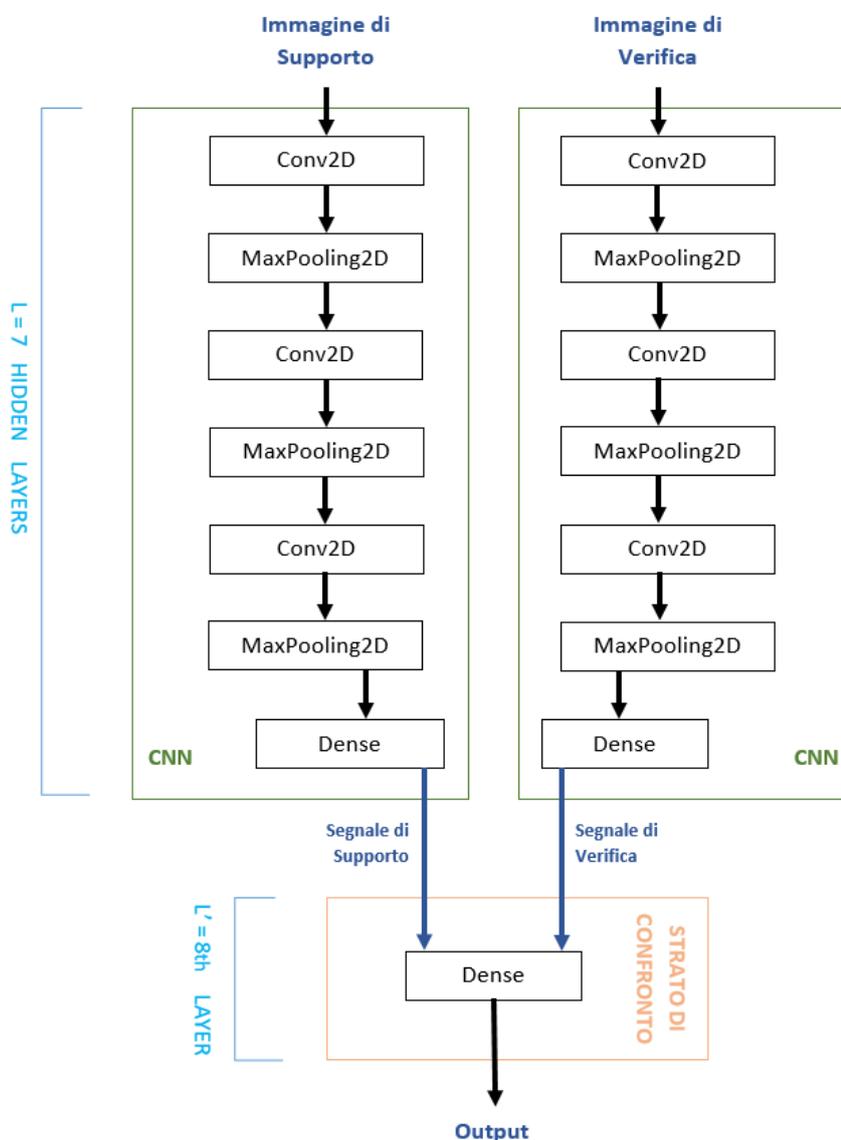


Figura 1 Rappresentazione concettuale dell'architettura di base della rete siamese utilizzata in questo lavoro.

⁸ Sono state considerate, nel corso della fase di sperimentazioni, ulteriori varianti. Per brevità e scarso contributo in termini di rilevanza ai fini delle considerazioni finali, non si riportano i risultati.

Per favorire il confronto tra risultati, l'architettura delle reti è invariata nei blocchi principali, mentre il numero di parametri è definito in funzione della risoluzione considerata. In particolare, a valle dell'acquisizione manuale delle immagini ad alta risoluzione della banconota (~2656x1467 pixel, 24 bit colore, secondo il modello RGB), sono stati prodotti tramite ri-campionamento i seguenti quattro insiemi di immagini, ciascuno dei quali è stato utilizzato, in sessioni distinte, per l'addestramento e il test della rete siamese, al fine di confrontare le prestazioni ottenute nei quattro scenari:

- Intera AR: Banconota intera, ad alta risoluzione (896x528 pixel);
- Intera MR: Banconota intera, a media risoluzione (448x264 pixel);
- Intera BR: Banconota intera, a bassa risoluzione (299x176 pixel);
- Bandiera: Ritaglio di parte della banconota⁹ ad alta risoluzione (224x352 pixel).

L'addestramento delle diverse reti siamesi avviene come detto attraverso l'algoritmo di *Stochastic Gradient Descent* (SGD), a partire da un set di parametri inizializzati secondo il metodo *Xavier*. Per favorire la convergenza ottimale del processo, si ricorre inoltre alla *batch normalization* e al *dropout*, nonché alla re-inizializzazione del valore dei parametri in caso di raggiungimento di un ottimo locale o di *overfitting* dell'insieme di *training* (figura 2). Si definiscono inoltre iper-parametri quali le funzioni di attivazione *Leaky ReLu* (Maas *et al.*, 2013) a partire da un insieme di immagini di calibrazione, escluse dal processo di addestramento.

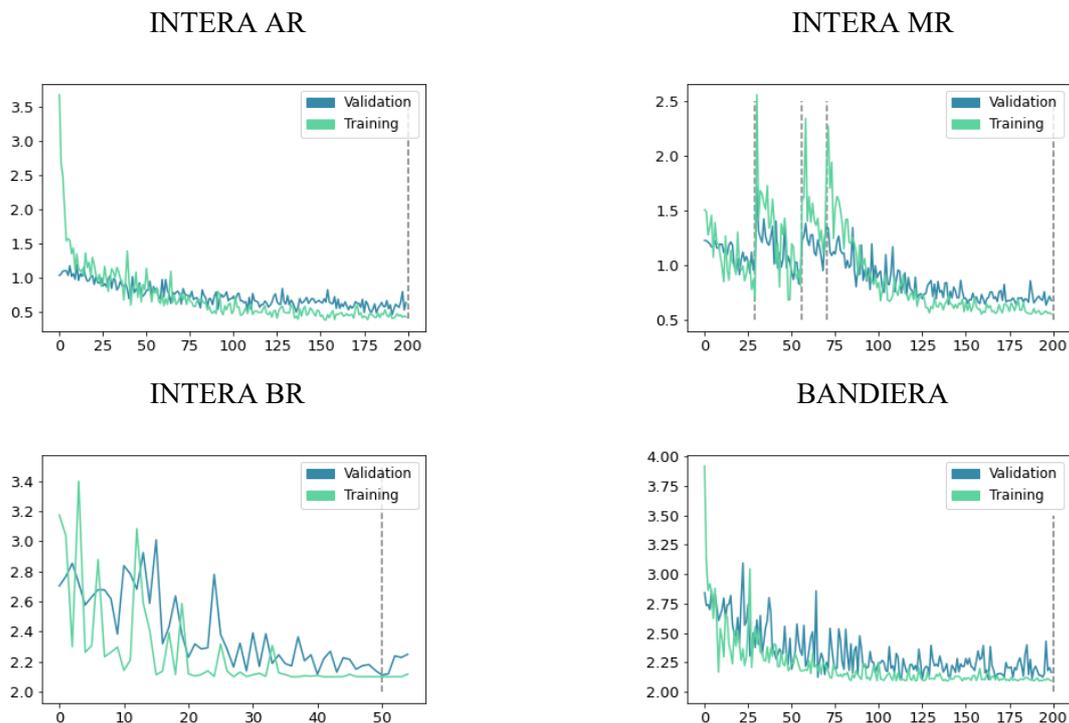


Figura 2 Processo di addestramento delle reti neurali siamesi per le diverse tipologie di input considerate: Intera AR, Intera MR, Intera BR e Bandiera. Le curve verdi (blu) rappresentano il valore della funzione di perdita in corrispondenza delle iterazioni della routine di ottimizzazione SGD nell'insieme di immagini di training (validazione). Le linee tratteggiate indicano la re-inizializzazione del processo di addestramento, in corrispondenza della convergenza presso un ottimo locale.

⁹ In linea con Jang *et al.* (2020).

L'elevata qualità della *performance* di algoritmi di *machine learning* deve spesso confrontarsi con una ridotta interpretabilità dei meccanismi interni appresi, nonché della spiegabilità delle logiche sottostanti. Le tecniche utili a supportare la comprensione, e quindi il controllo del corretto funzionamento di un modello analitico complesso, afferiscono al campo della cosiddetta *EXplainable Artificial Intelligence* (Molnar, 2020). Nel caso delle reti siamesi, la rappresentazione grafica dei segnali consente il confronto visivo tra la sintesi delle immagini di supporto e di verifica, per una valutazione qualitativa della similarità dei vettori estratti da ciascuna rete gemella. Tuttavia, la logica sottostante il processo di estrazione dei segnali è celata dall'elevato numero di parametri utilizzati per la proiezione ricorsiva del contenuto informativo delle immagini in input, in linea con la scarsa interpretabilità dei modelli di *deep learning* (Zhang *et al.*, 2019).

A differenza di altre tipologie di reti neurali, nelle CNN la trasmissione delle immagini di input da uno strato convoluzionale all'altro consente di estrarre dei tensori intermedi che possono essere a loro volta rappresentati come immagini, dette maschere di attenzione. In letteratura, il ricorso alle maschere di attenzione è utilizzato in genere durante l'addestramento del modello, ed è finalizzato al controllo esogeno del processo da parte dell'analista (Xu *et al.*, 2015). Intuitivamente, il tensore dei pesi dello strato convoluzionale l -esimo \mathbf{w}^l è combinato elemento per elemento con una maschera arbitraria \mathbf{a} atta a enfatizzare o meno zone specifiche dell'immagine, e trasformato come visto in precedenza secondo la funzione di attivazione g^l . Il tensore risultante, $\mathbf{w}^{l,a}$, è così caratterizzato da elementi tali per cui a valori più alti corrisponderà maggiore attenzione della rete per una particolare zona dell'immagine e viceversa.

Ai fini dell'applicazione considerata, i meccanismi di attenzione si dividono in *hard* e *soft*. L'attenzione di tipo *hard* corrisponde di fatto all'esclusione casuale o deterministica di parte della superficie (*cropping*). In questo caso \mathbf{a} assume valori binari: 0 o 1. I meccanismi di attenzione *hard* possono essere impiegati per indirizzare in modo esogeno il processo di apprendimento o per massimizzare l'esplorazione della superficie del tensore e definire delle traiettorie ottimali di attenzione da parte del modello (Ba *et al.*, 2014). Nella sua forma più semplice, l'attenzione di tipo *soft* ricorre invece ad una funzione di attivazione che comprime i valori dei pesi in un dato intervallo (Xu *et al.*, 2015), come ad esempio nell'intervallo $[0,1]$ nel caso della funzione *softmax*. I meccanismi di attenzione *soft* integrati sono utilizzati per garantire un controllo globale del processo di apprendimento¹⁰ (Gregor *et al.*, 2015; Kosiosek *et al.*, 2017).

L'applicazione in oggetto considera un uso alternativo di maschere di attenzione *soft* secondo un approccio semplificato, orientato alla spiegabilità del modello. Le maschere di attenzione sono attivate localmente, a valle dell'addestramento, per l'estrazione di immagini parziali lungo le CNN gemelle che rappresentino la logica adottata dal modello lungo le fasi dell'elaborazione. In particolare, per ciascuno strato convoluzionale il corrispondente tensore di pesi è proiettato nell'intervallo unitario $[0,1]$ secondo la funzione di attivazione *softmax*, che evidenzia il peso relativo di ciascun elemento. I pesi così trasformati sono quindi combinati con l'input dello strato convoluzionale. Ciò corrisponde alla sovrapposizione di una maschera di attenzione *soft*

¹⁰ In particolare, in corrispondenza di un dato output, come ad esempio la classificazione di un elemento sulla superficie dell'immagine, al modello è richiesto di identificare la regione di riferimento, operando una polarizzazione nel valore dei parametri corrispondenti.

all'immagine della banconota. L'intuizione è la seguente: lo strato l-esimo utilizza pesi i cui relativi valori nella maschera *soft* sono vicini ad 1 in corrispondenza di regioni considerate rilevanti ai fini dell'estrazione del segnale, e vicini a 0 viceversa.

4.2 Risultati

Ultimata la fase di apprendimento, la rete siamese classifica ogni nuova immagine utilizzando gli insiemi di immagini di supporto di cui per costruzione è nota l'appartenenza a una delle due categorie *bite* e *no-bite*, ovvero le categorie per i casi di presenza o assenza del difetto di stampa. L'assegnazione di una nuova immagine a una delle due classi avviene quindi a partire dal confronto tra l'immagine e i due insiemi di supporto.

La selezione delle immagini di supporto da utilizzare per il confronto costituisce un passaggio critico nell'uso delle reti siamesi in quanto la classificazione avviene a partire da coppie di osservazioni. La scelta sul metodo di selezione delle immagini di supporto ha certamente un impatto sulle prestazioni generali della rete; per analizzare questo impatto è stato costruito un insieme di immagini di *hold-out*, ovvero una porzione delle immagini dell'insieme di verifica, inizialmente separato da quello di *training* e utilizzato per la sola valutazione del metodo di selezione delle immagini di supporto, a garanzia della qualità dell'analisi. Si considerano quindi i seguenti approcci:

- 1) RND-1: ciascuna immagine di verifica è confrontata con una sola immagine di supporto per classe, selezionata casualmente, da considerarsi rappresentativa;
- 2) SSIM-0: ciascuna immagine di verifica è confrontata con una sola immagine di supporto per classe, selezionata in qualità di massimamente simile all'immagine di verifica tra quelle disponibili, secondo la *Structure Similarity Index Measure* (SSIM)¹¹;
- 3) SSIM-1: ciascuna immagine di verifica è confrontata con la seconda più simile tra le candidate di supporto disponibili per ciascuna classe. Rispetto all'approccio SSIM-0, l'esclusione dell'immagine maggiormente simile ha l'obiettivo di valutare le immagini al netto di caratteristiche che potrebbero risultare fuorvianti ai fini dell'esercizio di classificazione;
- 4) SSIM-k: ciascuna immagine di verifica è confrontata con k immagini di supporto per classe, selezionate secondo similarità (k=10 nell'analisi seguente).

Si riportano nella tabella 1 i risultati ottenuti utilizzando le quattro diverse risoluzioni descritte in precedenza per le immagini di verifica, e utilizzando i quattro diversi metodi di selezione per le immagini di supporto. L'insieme di test contiene 46 immagini, di cui 14 con difetto *bite* di diversa gravità¹². La *performance* sull'insieme di test è misurata a partire dal tasso di *Recall*¹³, vale a dire il tasso di immagini con difetto *bite* correttamente classificate, e dall'*F1-score*, media armonica del punteggio di *Recall* e di *Precision*, definito come tasso di banconote correttamente classificate come non difettose. Ai fini dell'applicazione, i falsi positivi, ovvero immagini di banconote erroneamente individuate come difettose, e l'*under-reporting* dei difetti, o falsi negativi, sono considerati ugualmente costosi.

¹¹ La SSIM è una misura diffusa per il confronto tra immagini e la valutazione della qualità delle stesse, introdotta da Wang *et al.* (2004). È derivata dalla composizione di tre elementi di confronto, relativi alla luminosità, al contrasto e alla struttura delle immagini.

¹² La classificazione è effettuata sulla base della documentazione prodotta dalla Banca Centrale Europea.

¹³ Dato il numero di banconote classificate come difettose dalla rete (*all positives*), il *Recall* è definito dalla proporzione di banconote correttamente classificate (*true positives*): $Recall = (true\ positives)/(all\ positives)$.

	Recall				F1-Score			
	RND-1	SSIM-0	SSIM-1	SSIM-k	SSIM-k	SSIM-0	SSIM-1	SSIM-k
Intera AR	-	.75	.62	-	-	.67	.55	-
Intera MR	1.	.87	.87	.5	.84	.82	.93	.67
Intera BR	-	.62	.62	-	-	.55	.62	-
Bandiera	-	-	.5	.98	-	-	-	-

Tabella 1 Risultati relativi all’esercizio di classificazione binaria delle immagini nell’insieme di test, per i casi di banconota intera ad alta (AR), media (MR) e bassa (BR) risoluzione, e per il caso di dettaglio della bandiera europea. Le *performance* inferiori al valore soglia 0.5, associato ad un classificatore casuale, sono indicate con “-“.

Si rileva come, limitatamente all’insieme di verifica considerato, l’utilizzo dell’immagine intera a media risoluzione (Intera MR) consenta di ottenere maggiore accuratezza nella classificazione delle osservazioni in *bite* e *no-bite*. In particolare, il metodo di selezione SSIM-1 risulta essere maggiormente performante nell’assegnazione della banconota all’una o all’altra categoria.

Le reti siamesi sono caratterizzate dall’abilità nell’identificare le caratteristiche critiche ai fini della proiezione dell’input in forma di segnale e, quindi, la corretta assegnazione dell’immagine di verifica alla classe di appartenenza, in condizioni di elevata variabilità. Quando questa risulti essere eccessiva, come nel caso dell’analisi a partire dalle immagini a maggiore risoluzione (Intera AR) la granularità in termini di pixel diventa fuorviante per la separabilità dei segnali ottenuti per le immagini di verifica e supporto. Ciò è riscontrabile anche a partire dall’ispezione visiva dei segnali estratti dalla rete siamese, riportati nel riquadro in alto a sinistra della figura 3, in cui si nota come le differenze tra i segnali ottenuti dalle immagini di supporto alle due categorie siano tra loro più simili, se confrontate con il caso a media risoluzione (in alto a destra). In maniera speculare, riducendo eccessivamente la risoluzione dell’immagine, la dimensione assoluta del difetto *bite*, quella relativa del campo corrispondente alla bandiera e la complessiva variabilità risultano essere attenuate fino al punto da non permettere alla rete siamese di apprendere un criterio affidabile per il loro riconoscimento (in basso a sinistra). Nel caso dell’analisi a partire da immagini circoscritte all’elemento della bandiera, il combinato disposto tra la riduzione della variabilità complessiva (di secondo livello) e l’aumentato peso relativo della regione di pixel corrispondenti al difetto, caratterizzato da un’alta variabilità (di primo livello), rendono inadeguato un approccio basato su apprendimento *one-shot*. La *performance*, misurata in termini quantitativi, e l’ispezione grafica dei segnali ottenuti (in basso a destra) riflettono infatti tale inadeguatezza.

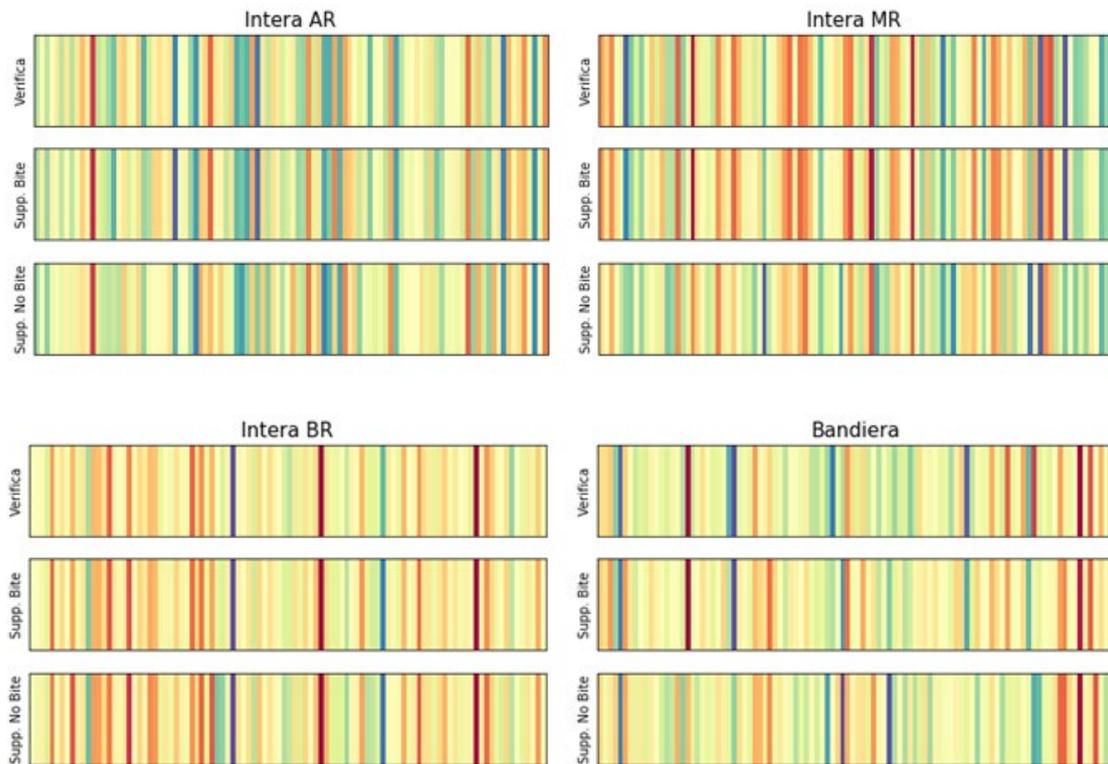


Figura 3 Confronto tra segnali per le immagini di verifica e di supporto. Le immagini riportano i segnali ottenuti per coppie di immagini di tipo Intera AR, Intera MR, Intera BR, e Bandiera. Le immagini di supporto sono state selezionate utilizzando il metodo SSIM-1. Per ciascun riquadro, il primo segnale a partire dall'alto è ottenuto dalla immagine di verifica con difetto *bite*, da confrontare, rispettivamente, con le immagini di supporto di tipo *bite* e *no-bite*.

La figura 4 illustra un esempio di applicazione del meccanismo di attenzione operato dalle componenti convoluzionali delle reti siamesi, ottenuto per la stessa immagine con difetto *bite* considerata nei formati Intera AR, MR, BR, e Bandiera. Nelle diverse immagini, a gradazioni di colore più chiare corrisponde una maggiore attenzione, o concentrazione, conferita dalla rete neurale nei tre strati convoluzionali interni alla CNN. Si può notare come, in assenza di indicazioni fornite alla rete in fase di costruzione della stessa, sia l'architettura per l'analisi di immagini ad alta risoluzione sia quella per le immagini a media risoluzione siano in grado di indentificare autonomamente l'area della superficie critica ai fini della rilevazione del difetto *bite* e la conseguente assegnazione a una classe o all'altra. Tale processo virtuoso non sembra innescarsi, se non in via approssimata, nel caso della rete siamese calibrata a partire dall'insieme di *training* a bassa risoluzione. Infine, si produce la reazione speculare nel caso relativo alla sola bandiera: la rete siamese sembra "concentrarsi" sull'intera area dell'immagine, rendendo così il processo di sintesi dell'informazione in input eccessivamente suscettibile a variazioni di lieve entità.

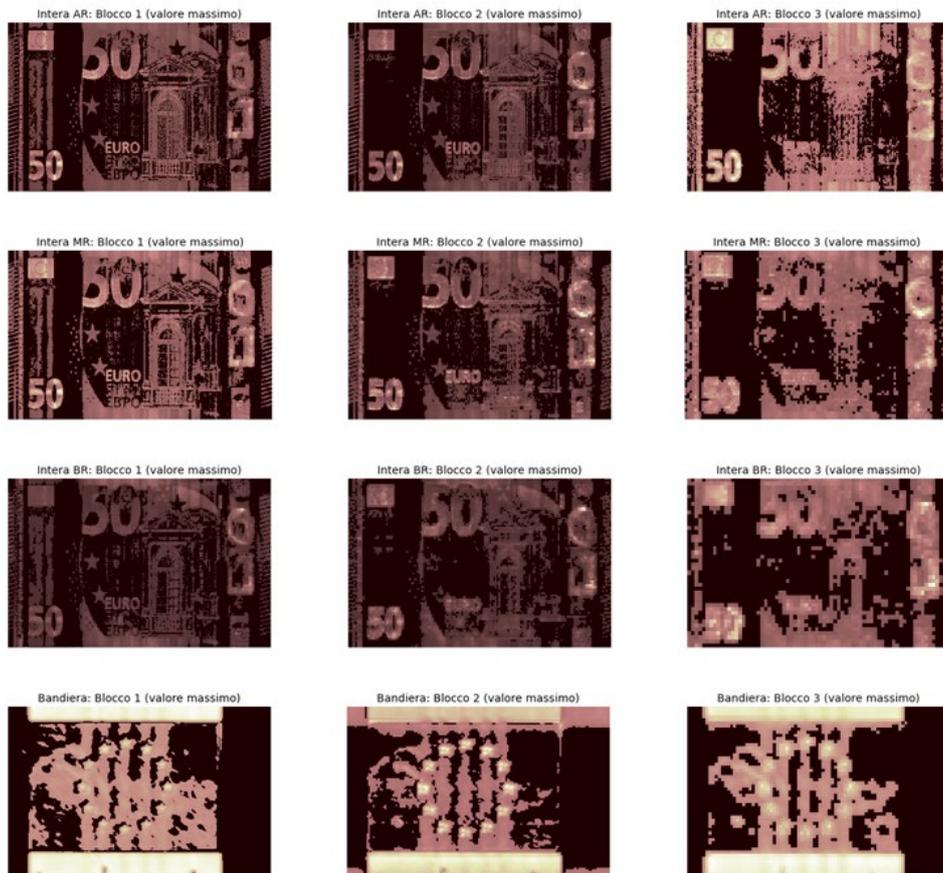


Figura 4 Rappresentazione grafica del meccanismo di attenzione operato dai tre blocchi della componente convoluzionale delle reti siamesi per l’immagine di *training* con difetto *bite* nei formati Intera AR, Intera MR, Intera BR, e Bandiera.

5. Conclusioni e possibili estensioni

Il presente lavoro riporta i principali risultati ottenuti dallo studio esplorativo sull’uso di reti neurali per il riconoscimento e la classificazione “intelligente” dei difetti delle banconote nell’ambito dei controlli di qualità. La complessità del problema, dovuta principalmente all’alto grado di naturale variabilità del prodotto e alla variabilità di tipologia e posizione dei difetti, ha imposto un approccio molto stringente nella selezione degli scenari di sperimentazione. Ciò ha ristretto l’analisi delle potenzialità di tali strumenti ad una specifica tipologia di difetti (*bite*), localizzata su una prestabilita porzione (Bandiera) di un unico taglio di banconote (50€ della seconda serie).

I risultati supportano la rilevanza di questi strumenti per l’identificazione di eventuali difetti di stampa. In particolare, il caso di banconota intera a media risoluzione (Intera MR) consente di bilanciare le esigenze legate all’acquisizione delle immagini, l’efficienza di calcolo e la qualità della *performance* dei modelli. Il lavoro di analisi da svolgere per poterne attestare l’effettiva potenzialità di impiego alternativo, ovvero cooperativo, alle attuali tecniche qualitative non automatiche è da considerarsi di natura preliminare. Ulteriori attività di studio e sperimentazione si propongono di estendere l’analisi a categorie di difetti diverse dal *bite*, avvalendosi di tecniche utili ad adattare a un nuovo compito un modello precedentemente addestrato (c.d. *transfer learning*), mantenendo un

output di classificazione binario. Infine, l'implementazione di un sistema per la classificazione di difetti di stampa multipli potrebbe avvalersi di tecniche di c.d. *ensemble learning* per combinare e aggregare i risultati prodotti da un'architettura multi-rete.

Le eventuali estensioni considerate sarebbero comunque circoscritte da una scelta a-priori che ne mantiene il dominio di analisi e ricerca all'interno di algoritmi di tipo *one-shot learning*, e nello specifico utilizzando reti siamesi. Questa scelta è suggerita anche dalla letteratura prodotta dalla comunità scientifica per problemi a vario titolo riconducibili a quanto presentato in queste pagine.

Per concludere, la *performance* dei modelli è stata valutata su un ulteriore insieme di immagini ottenuto attraverso un processo meno raffinato di scansione digitale delle banconote. I risultati ottenuti si sono rivelati in linea con quanto osservato nel campione di test, e sono dettagliati in Appendice. Sono state inoltre incluse in questa ulteriore analisi 11 immagini con difetto di tipo *bite-L*, un difetto di dimensioni ridotte che, sebbene non sia una imperfezione che imponga di scartare un intero lotto di produzione di banconote, è comunque segnalata dall'operatore così da fornire un *feedback* di qualità al processo produttivo, al fine di evitare una prolungata presenza di tali imperfezioni, o la presenza di difetti di maggiore gravità. Nel caso specifico, l'inclusione di tale gruppo di difetti ha carattere di controllo ed è volta a valutare l'abilità del modello di classificare tali immagini come difettose.

Riferimenti bibliografici

- Ali, T., Jan, S., Alkhodre, A., Nauman, M., Amin, M., & Siddiqui, M.S. (2019). DeepMoney: counterfeit money detection using generative adversarial networks. *PeerJ Computer Science*, 5, e216.
- Aseffa, D.T., Kalla, H., & Mishra, S. (2022). Ethiopian banknote recognition using convolutional neural network and its prototype development using embedded platform. *Journal of Sensors*, 2022, 1-18.
- Ba, J.L., Mnih, V., & Kavukcuoglu, K. (2014). Multiple object recognition with visual attention.
- Boureau, Y.L., Ponce, J., & LeCun, Y. (2010). A theoretical analysis of feature pooling in visual recognition. In *Proceedings of the 27th international conference on machine learning (ICML-10)* (pp. 111-118).
- Darade, S.R., & Gidveer, G.R. (2016). Automatic recognition of fake Indian currency note. In *2016 international conference on Electrical Power and Energy Systems (ICEPES)* (pp. 290-294). IEEE.
- Daraee, F., & Mozaffari, S. (2010). Eroded money notes recognition using wavelet transform. In *2010 6th Iranian Conference on Machine Vision and Image Processing* (pp. 1-5). IEEE.
- Desai, S., Rajadhyaksha, A., Shetty, A., & Gharat, S. (2021). CNN based counterfeit Indian currency recognition using generative adversarial network. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)* (pp. 626-631). IEEE.
- Domínguez, F.R. (2020). Supermasks and a Good Initialization Are All You Need (*Doctoral dissertation, Pontificia Universidad Católica de Chile (Chile)*).
- Fei-Fei, Li, Rob Fergus e Pietro Perona (2006). One-shot learning of object categories. *IEEE transactions on pattern analysis and machine intelligence* 28.4: 594-611.
- Feng, B.Y., Ren, M., Zhang, X.Y., & Suen, C.Y. (2014). Part-based high accuracy recognition of serial numbers in bank notes. In *Artificial Neural Networks in Pattern Recognition: 6th IAPR TC 3 International Workshop, ANNPR 2014, Montreal, QC, Canada, October 6-8, 2014. Proceedings 6* (pp. 204-215). Springer International Publishing.
- García-Lamont, F., Cervantes, J., López, A., & Rodríguez, L. (2013). Classification of Mexican paper currency denomination by extracting their discriminative colors. In *Advances in Soft Computing and Its Applications: 12th Mexican International Conference on Artificial Intelligence, MICAI 2013, Mexico City, Mexico, November 24-30, 2013, Proceedings, Part II 12* (pp. 403-412). Springer Berlin Heidelberg.
- Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (pp. 249-256). JMLR Workshop and Conference Proceedings.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. *MIT press*.

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144.
- Gregor, K., Danihelka, I., Graves, A., Rezende, D., & Wierstra, D. (2015). Draw: A recurrent neural network for image generation. In *International conference on machine learning* (pp. 1462-1471). PMLR.
- Grijalva, F., Rodriguez, J.C., Larco, J., & Orozco, L. (2010). Smartphone recognition of the US banknotes' denomination, for visually impaired people. In *2010 IEEE ANDESCON* (pp. 1-6). IEEE.
- Guo, Z.H., & Li, S.H. (2006). Feed-Forward neural network using SARPROP algorithm and its application in radar target recognition. In *Advances in Neural Networks-ISNN 2006: Third International Symposium on Neural Networks, Chengdu, China, May 28-June 1, 2006, Proceedings, Part II 3* (pp. 369-374). Springer Berlin Heidelberg.
- Hasanuzzaman, F.M., Yang, X., & Tian, Y. (2011). Robust and effective component-based banknote recognition by SURF features. In *2011 20th Annual Wireless and Optical Communications Conference (WOCC)* (pp. 1-6). IEEE.
- Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R.R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). pmlr.
- Jadhav, M., kumar Sharma, Y., & Bhandari, G.M. (2019). Currency identification and forged banknote detection using deep learning. In *2019 International Conference on Innovative Trends and Advances in Engineering and Technology (ICITAET)* (pp. 178-183). IEEE.
- Jang, U., Suh, K.H., & Lee, E.C. (2020). Low-quality banknote serial number recognition based on deep neural network. *Journal of Information Processing Systems*, 16(1), 224-237.
- Kamble, K., Bhansali, A., Satalgaonkar, P., & Alagundgi, S. (2019). Counterfeit currency detection using deep convolutional neural network. In *2019 IEEE Pune Section International Conference (PuneCon)* (pp. 1-4). IEEE.
- Ke, W., Huiqin, W., Yue, S., Li, M., & Fengyan, Q. (2016). Banknote image defect recognition method based on convolution neural network. *International Journal of Security and Its Applications*, 10(6), 269-280.
- Khemiri, W., Tarifa, A., Jaafar, W., & Abderrazak, J.B. (2022). Towards a Hybrid Variant of GAN For Counterfeit Money Detection SSQUGAN: A Semi-Supervised Quadrupled GAN.
- Koch, G., Zemel, R., & Salakhutdinov, R. (2015). Siamese neural networks for one-shot image recognition. In *ICML deep learning workshop* (Vol. 2, No. 1).
- Kosiorok, A., Bewley, A., & Posner, I. (2017). Hierarchical attentive recurrent tracking. *Advances in neural information processing systems*, 30.
- Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.

- Laavanya, M., & Vijayaraghavan, V. (2019). Real time fake currency note detection using deep learning. *Int. J. Eng. Adv. Technol.(IJEAT)*, 9.
- LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*, 2.
- Lee, J. W., Hong, H.G., Kim, K.W., & Park, K.R. (2017). A survey on banknote recognition methods by various sensors. *Sensors*, 17(2), 313.
- Linkon, A.H.M., Labib, M.M., Bappy, F.H., Sarker, S., Jannat, M.E., & Islam, M.S. (2020). Deep Learning Approach Combining Lightweight CNN Architecture with Transfer Learning: An Automatic Approach for the Detection and Recognition of Bangladeshi Banknotes. In *2020 11th International Conference on Electrical and Computer Engineering (ICECE)* (pp. 214-217). IEEE.
- Liu, L., Ye, Y.T., Xie, Y., & Pu, L. (2010). Serial number extracting and recognizing applied in paper currency sorting system based on RBF Network. In *2010 international conference on computational intelligence and software engineering* (pp. 1-4). IEEE.
- Maas, A.L., Hannun, A.Y., & Ng, A.Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml* (Vol. 30, No. 1, p. 3).
- Manikandan, K., & Sumithra, T. (2015). Currency recognition in mobile application for visually challenged. *Discovery*, 30, 245-248.
- Molnar, C. (2020). Interpretable machine learning. *Lulu.com*.
- Mousavi, S.A., Meghdadi, M., Hanifeloo, Z., Sumari, P., & Arshad, M.R.M. (2015). Old and worn banknote detection using sparse representation and neural networks. *Indian J. Sci. Technol*, 8, 913-918.
- Ochal, M., Patacchiola, M., Storkey, A., Vazquez, J., & Wang, S. (2021). Few-shot learning with class imbalance. *arXiv preprint arXiv:2101.02523*.
- Palatucci, M., Pomerleau, D., Hinton, G.E., & Mitchell, T.M. (2009). Zero-shot learning with semantic output codes. *Advances in neural information processing systems*, 22.
- Park, C., Cho, S.W., Baek, N.R., Choi, J., & Park, K.R. (2020). Deep feature-based three-stage detection of banknotes and coins for assisting visually impaired people. *IEEE Access*, 8, 184598-184613.
- Pham, T.D., Lee, D.E., & Park, K.R. (2017). Multi-national banknote classification based on visible-light line sensor and convolutional neural network. *Sensors*, 17(7), 1595.
- Qian, N. (1999). On the momentum term in gradient descent learning algorithms. *Neural networks*, 12(1), 145-151.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

- Sawant, V.M., Tupe, R. D., Tawade, A.S., & Bhalerao, S.M. (2022). Fake Currency Identification System Using CNN. *International Journal of Wireless Network Security*, 8(1), 18-23.
- Sharma, B., & Kaur, A. (2012). Recognition of Indian paper currency based on LBP. *International Journal of Computer Applications*, 59(1).
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sun, B., & Li, J. (2008). The recognition of new and old banknotes based on SVM. In *2008 Second International Symposium on Intelligent Information Technology Application* (Vol. 2, pp. 95-98). IEEE.
- Suresh, I.A., & Narwade, P.P. (2016). Indian currency recognition and verification using image processing. *International Research Journal of Engineering and Technology (IRJET)*, 3(6), 87-91.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
- Treadgold, N.K., & Gedeon, T.D. (1996). The SARPROP algorithm, a simulated annealing enhancement to resilient back propagation. In *Proceedings International Panel Conference on Soft and Intelligent Computing* (pp. 293-298).
- Veeramsetty, V., Singal, G., & Badal, T. (2020). Coinnet: platform independent application to recognize Indian currency notes using deep learning techniques. *Multimedia Tools and Applications*, 79(31-32), 22569-22594.
- Wang, J., & Perez, L. (2017). The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Networks Vis. Recognit*, 11, 1-8.
- Wenhong, L., Wenjuan, T., Xiyan, C., & Zhen, G. (2010). Application of support vector machine (SVM) on serial number identification of RMB. In *2010 8th World Congress on Intelligent Control and Automation* (pp. 6262-6266). IEEE.
- Wijnhoven, R.G., & de With, P.H.N. (2010). Fast training of object detection using stochastic gradient descent. In *2010 20th International conference on pattern recognition* (pp. 424-427). IEEE.
- Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R. & Bengio, Y. (2015). Show, attend and tell: Neural image caption generation with visual attention. In *International conference on machine learning* (pp. 2048-2057). PMLR.
- Yan, Z., Zhang, H., Piramuthu, R., Jagadeesh, V., DeCoste, D., Di, W., & Yu, Y. (2015). HD-CNN: hierarchical deep convolutional neural networks for large scale visual recognition. In *Proceedings of the IEEE international conference on computer vision* (pp. 2740-2748).
- Yu, K., Xu, W., & Gong, Y. (2008). Deep learning with kernel regularization for visual recognition. *Advances in neural information processing systems*, 21.

Wang Z., Bovik, A.C., Sheikh H.R., Simoncelli, E.P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* 13.4: 600-612.

Zeiler, M.D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13* (pp. 818-833). Springer International Publishing.

Zhang, Q., Yan, W.Q., & Kankanhalli, M. (2019). Overview of currency recognition using deep learning. *Journal of Banking and Financial Technology*, 3, 59-69.

Appendice: Estensione a difetti *bite-L*

La *performance* dei modelli è stata valutata su un ulteriore insieme di immagini, ottenuto attraverso un processo meno raffinato di scansione digitale delle banconote (~675x388 pixel, 24 bit colore RGB, vale a dire ridotta risoluzione in fase di acquisizione, oltre a una maggiore variabilità nella centratura delle banconote rispetto al riquadro, colorazione più compressa e saturata basata su una *palette* meno ricca di sfumature). I risultati ottenuti si sono rivelati in linea con quanto osservato nel campione di test, in particolare nell'identificare correttamente il difetto, mentre si è registrata una tendenza più conservativa nell'escluderne la presenza. Ad esempio, per il modello a media risoluzione valutato secondo il meccanismo SSIM-1 si è ottenuto lo stesso valore della statistica *Recall* (87.5%), a fronte di un peggioramento della capacità complessiva del modello di valutare correttamente le banconote prive di difetti di tipo *bite* (0.61 *F1-Score*, -34.4%).

Sono state inoltre incluse nell'analisi 11 immagini con difetto *bite-L*, un difetto di dimensioni ridotte che, sebbene non sia una imperfezione che imponga di scartare un intero lotto di produzione di banconote, è comunque segnalata dall'operatore così da fornire un *feedback* di qualità al processo produttivo, al fine di evitare una prolungata presenza di tali imperfezioni o la presenza di difetti di maggiore gravità. Nel caso specifico, l'inclusione di tale gruppo di difetti ha carattere di controllo ed è volta a valutare l'abilità del modello di classificare tali immagini come difettose.

La *performance* delle reti siamesi è stata misurata per i diversi formati delle immagini in input ad alta e media risoluzione (Intera AR e Intera MR, introdotte nelle sezioni precedenti). Nel primo caso, la previsione risulta essere meno polarizzata rispetto al secondo. Con riferimento alla classificazione basata su SSIM-1, la rete siamese per le osservazioni di tipo Intera AR classifica 7 immagini su 11 alla classe *bite-L*, di cui 4 con probabilità tra 0,5 e 0,52.

Nel caso della rete siamese addestrata a partire da immagini di banconote a media risoluzione, 7 osservazioni sono classificate correttamente, con un punteggio di probabilità in media più elevato. Le reti siamesi per banconote a bassa risoluzione e con analisi limitata all'area della bandiera, applicando il metodo di selezione SSIM-1 delle immagini di supporto, classificano correttamente 6 e 5 osservazioni, rispettivamente.

Come ulteriore analisi, è stata considerata una rete siamese per la classificazione multi-etichetta di banconote secondo le categorie di i) *bite*, ii) *bite-L* (imperfezioni) e iii) *no-bite*, per i formati Intera MR e Bandiera. La classificazione multi-etichetta ha prodotto dei risultati preliminari soltanto in parte soddisfacenti: le reti siamesi sono riuscite ad apprendere un sistema efficace per distinguere i difetti *bite* da quelli di tipo *bite-L*, a svantaggio dell'abilità nel classificare correttamente le osservazioni *no-bite*, ossia si è osservato un aumento dei falsi positivi. La figura A.1 riporta i segnali ottenuti per la stessa immagine di tipo *bite* nei casi Intera MR e Bandiera: si osserva come la proiezione ottenuta a supporto della classe *bite-L* sia nel primo caso molto simile, a un'ispezione visiva, a quella *bite*, e nel secondo a quella *no-bite*.

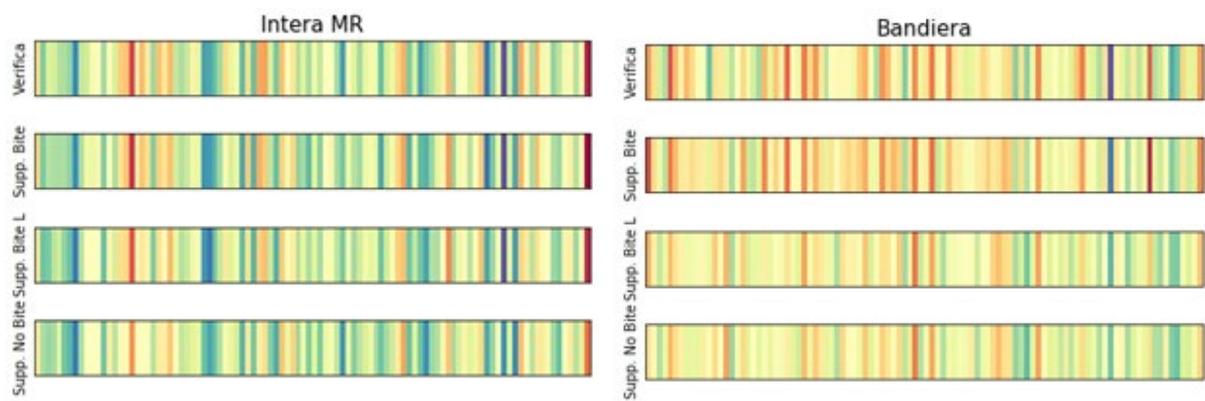


Figura A.1 Confronto tra segnali per le immagini di verifica e di supporto per le classi *bite*, *bite-L* e *no-bite*. I segnali fanno riferimento ad immagini a media risoluzione (sinistra) e ristrette all'area della bandiera (destra). Il metodo di selezione delle immagini di supporto è SSIM-1. Dall'alto verso il basso: segnale ottenuto dall'immagine di verifica con difetto *bite*, e segnali di supporto per le classi *bite*, *bite-L* e *no-bite*.

PUBBLICAZIONI DELLA COLLANA **MERCATI, INFRASTRUTTURE, SISTEMI DI PAGAMENTO**

- n. 1 TIPS - TARGET Instant Payment Settlement - Il sistema europeo per il regolamento dei pagamenti istantanei, *di Massimiliano Renzetti, Serena Bernardini, Giuseppe Marino, Luca Mibelli, Laura Ricciardi, Giovanni M. Sabelli* (QUESTIONI ISTITUZIONALI)
- n. 2 Real-Time Gross Settlement systems: breaking the wall of scalability and high availability, *di Mauro Arcese, Domenico Di Giulio, Vitangelo Lasorella* (APPROFONDIMENTI)
- n. 3 Green Bonds: the Sovereign Issuers' Perspective, *di Raffaele Doronzo, Vittorio Siracusa, Stefano Antonelli* (APPROFONDIMENTI)
- n. 4 T2S - TARGET2-Securities - La piattaforma paneuropea per il regolamento dei titoli in base monetaria, *di Cristina Mastropasqua, Alessandro Intonti, Michael Jennings, Clara Mandolini, Massimo Maniero, Stefano Vespucci, Diego Toma* (QUESTIONI ISTITUZIONALI)
- n. 5 The carbon footprint of the Target Instant Payment Settlement (TIPS) system: a comparative analysis with Bitcoin and other infrastructures, *di Pietro Tiberi* (APPROFONDIMENTI)
- n. 6 Proposal for a common categorisation of IT incidents, *di Autorité de Contrôle Prudentiel et de Résolution, Banca d'Italia, Commissione Nazionale per le Società e la Borsa, Deutsche Bundesbank, European Central Bank, Federal Reserve Board, Financial Conduct Authority, Ministero dell'Economia e delle Finanze, Prudential Regulation Authority, U.S. Treasury* (QUESTIONI ISTITUZIONALI)
- n. 7 Inside the black box: tools for understanding cash circulation, *di Luca Baldo, Elisa Bonifacio, Marco Brandi, Michelina Lo Russo, Gianluca Maddaloni, Andrea Nobili, Giorgia Rocco, Gabriele Sene, Massimo Valentini* (APPROFONDIMENTI)
- n. 8 L'impatto della pandemia sull'uso degli strumenti di pagamento in Italia, *di Guerino Ardizzi, Alessandro Gambini, Andrea Nobili, Emanuele Pimpini, Giorgia Rocco* (APPROFONDIMENTI)
- n. 9 TARGET2 - Il sistema europeo per il regolamento dei pagamenti di importo rilevante, *di Paolo Bramini, Matteo Coletti, Francesco Di Stasio, Pierfrancesco Molina, Vittorio Schina, Massimo Valentini* (QUESTIONI ISTITUZIONALI)
- n. 10 A digital euro: a contribution to the discussion on technical design choices, *di Emanuele Urbinati, Alessia Belsito, Daniele Cani, Angela Caporini, Marco Capotosto, Simone Folino, Giuseppe Galano, Giancarlo Goretti, Gabriele Marcelli, Pietro Tiberi, Alessia Vita* (QUESTIONI ISTITUZIONALI)
- n. 11 From SMP to PEPP: A Further Look at the Risk Endogeneity of the Central Bank, *di Marco Fruzzetti, Giulio Gariano, Gerardo Palazzo, Antonio Scalia* (APPROFONDIMENTI)
- n. 12 Le TLTRO e la disponibilità di garanzie in Italia, *di Annino Agnes, Paola Antilici, Gianluca Mosconi* (APPROFONDIMENTI)
- n. 13 Overview of central banks' in-house credit assessment systems in the euro area, *di Laura Auria, Markus Bingmer, Carlos Mateo Caicedo Graciano, Clémence Charavel, Sergio Gavilá, Alessandra Iannamorelli, Aviram Levy, Alfredo Maldonado, Florian Resch, Anna Maria Rossi, Stephan Sauer* (QUESTIONI ISTITUZIONALI)
- n. 14 L'allocazione strategica e la sostenibilità degli investimenti della banca centrale, *di Davide Di Zio, Marco Fanari, Simone Letta, Tommaso Perez, Giovanni Secondin* (APPROFONDIMENTI)
- n. 15 Climate and environmental risks: measuring the exposure of investments, *di Ivan Faiella, Enrico Bernardini, Johnny Di Giampaolo, Marco Fruzzetti, Simone Letta, Raffaele Loffredo, Davide Nasti* (APPROFONDIMENTI)

- n. 16 Cross-Currency Settlement of Instant Payments in a Multi-Currency Clearing and Settlement Mechanism, *di Massimiliano Renzetti, Fabrizio Dinacci, Ann Börestam* (APPROFONDIMENTI)
- n. 17 Quale futuro per i benchmark del mercato monetario in euro?, *di Daniela Della Gatta* (QUESTIONI ISTITUZIONALI)
- n. 18 Cyber resilience per la continuità di servizio del sistema finanziario, *di Boris Giannetto, Antonino Fazio* (QUESTIONI ISTITUZIONALI)
- n. 19 Cross-Currency Settlement of Instant Payments in a Cross-Platform Context: a Proof of Concept, *di Massimiliano Renzetti, Andrea Dimartina, Riccardo Mancini, Giovanni Sabelli, Francesco Di Stasio, Carlo Palmers, Faisal Alhijawi, Erol Kaya, Christophe Piccarelle, Stuart Butler, Jwallant Vasani, Giancarlo Esposito, Alberto Tiberino, Manfredi Caracausi* (APPROFONDIMENTI)
- n. 20 Flash crashes on sovereign bond markets – EU evidence, *di Antoine Bouveret, Martin Haferkorn, Gaetano Marseglia, Onofrio Panzarino* (APPROFONDIMENTI)
- n. 21 Report on the payment attitudes of consumers in Italy: results from ECB surveys, *di Gabriele Coletti, Alberto Di Iorio, Emanuele Pimpini, Giorgia Rocco* (QUESTIONI ISTITUZIONALI)
- n. 22 When financial innovation and sustainable finance meet: Sustainability-Linked Bonds, *di Paola Antilici, Gianluca Mosconi, Luigi Russo* (QUESTIONI ISTITUZIONALI)
- n. 23 Business models and pricing strategies in the market for ATM withdrawals, *di Guerino Ardizzi, Massimiliano Cologgi* (APPROFONDIMENTI)
- n. 24 Press news and social media in credit risk assessment: the experience of Banca d'Italia's In-house Credit Assessment System, *di Giulio Gariano, Gianluca Viggiano* (APPROFONDIMENTI)
- n. 25 The bonfire of banknotes, *di Michele Manna* (APPROFONDIMENTI)
- n. 26 Integrating DLTs with market infrastructures: analysis and proof-of-concept for secure DvP between TIPS and DLT platforms, *di Rosario La Rocca, Riccardo Mancini, Marco Benedetti, Matteo Caruso, Stefano Cossu, Giuseppe Galano, Simone Mancini, Gabriele Marcelli, Piero Martella, Matteo Nardelli, Ciro Oliviero* (APPROFONDIMENTI)
- n. 27 Uso statistico e previsivo delle transazioni elettroniche di pagamento: la collaborazione Banca d'Italia-Istat, *di Guerino Ardizzi e Alessandra Righi* (QUESTIONI ISTITUZIONALI)
- n. 28 TIPS: a zero-downtime platform powered by automation, *di Gianluca Caricato, Marco Capotosto, Silvio Orsini, Pietro Tiberi* (APPROFONDIMENTI)
- n. 29 TARGET2 analytical tools for regulatory compliance, *di Marc Glowka, Alexander Müller, Livia Polo Friz, Sara Testi, Massimo Valentini, Stefano Vespucci* (QUESTIONI ISTITUZIONALI)
- n. 30 The security of retail payment instruments: evidence from supervisory data, *di Massimiliano Cologgi* (APPROFONDIMENTI)
- n. 31 Open Banking in the payment system: infrastructural evolution, innovation and security, supervisory and oversight practices, *di Roberto Pellitteri, Ravenio Parrini, Carlo Cafarotti, Benedetto Andrea De Vendictis* (QUESTIONI ISTITUZIONALI)
- n. 32 Banks' liquidity transformation rate: determinants and impact on lending, *di Raffaele Lenzi, Stefano Nobili, Filippo Perazzoli, Rosario Romeo* (APPROFONDIMENTI)
- n. 33 Investor behavior under market stress: evidence from the Italian sovereign bond market, *di Onofrio Panzarino* (APPROFONDIMENTI)