

ALMA MATER STUDIORUM – UNIVERSITA' DI BOLOGNA

CAMPUS DI CESENA

SCUOLA DI SCIENZE

CORSO DI LAUREA IN SCIENZE DELL'INFORMAZIONE

**Tecniche di riconoscimento del volto in
presenza di makeup**

Relazione finale in

Basi di Dati

Relatore
Dott.ssa Annalisa Franco

Presentata da
Andrea Farneti

Sessione III

Anno Accademico 2014-2015

Indice

Introduzione.....	1
Capitolo 1: Riconoscimento del volto	5
1.1 Individuazione del volto	7
1.1.1 Skin Segmentation.....	8
1.1.2 Metodo Viola-Jones.....	9
1.2 Ottimizzazione	11
1.3 Da immagine a vettore	14
1.3.1 Estrazione delle feature	15
1.3.2 Dimensionality Reduction	17
1.4 Riconoscimento.....	18
1.4.1 Support Vector Machine	19
1.4.2 Eigenfaces.....	19
1.4.3 Hidden Markov Model	20
1.4.4 Adaboost	22
Capitolo 2: Makeup.....	23
2.1 Applicazione del makeup	25
2.2 Software per il makeup virtuale	30
2.2.1 Mary Kay Virtual Makeover.....	31
2.2.2 TaaZ Virtual Makeover	33
2.3 Database.....	35
2.3.1 Youtube Makeup database	36
2.3.2 Virtual Makeup database	38
2.3.3 Makeup In the Wild database	39
2.3.4 Facial Cosmetics Database	40
2.3.5 FAcE Makeup database	42
2.4 Impatto del makeup sul riconoscimento del volto	42
2.4.1 Esperimento 1 su YMU	43
2.4.2 Esperimento 2 su VMU.....	46
2.4.3 Esperimento 3 su FCD	49

2.4.4 Conclusioni	52
Capitolo 3: Riconoscimento del volto in presenza di makeup.....	55
3.1 Algoritmi in grado di rilevare il trucco	56
3.1.1 Algoritmo basato su colore, texture e forma	56
3.1.2 Rilevamento del makeup tramite l'uso del Local Fisher Discriminant Analysis	61
3.1.3 Algoritmo basato sull'analisi di forma e texture (LGBP + HOG).....	66
3.2 Algoritmi in grado di rimuovere il trucco	72
3.2.1 Algoritmo basato sull'algebra di Clifford.....	72
3.3 Algoritmi di confronto robusti al problema del makeup.....	77
3.3.1 Makeup detector applicato al riconoscimento del volto	77
3.3.2 Riconoscimento facciale (LBP + DT-CWT)	79
3.4 VeriLook 5.3/MegaMatcher 4.4 Algorithm	82
3.4.1 Requisiti di qualità delle immagini	83
3.4.2 Configurazione del software	84
3.4.3 Esperimenti	89
Conclusioni.....	97
Ringraziamenti	99
Bibliografia e Sitografia.....	101

Introduzione

Il riconoscimento del volto inteso come rilevamento automatico da parte di un sistema completamente automatizzato è un concetto relativamente recente. Sviluppato da Woody Bledsoe, Helen Chan Wolf, and Charles Bisson nel 1960 [4], il primo sistema semi automatico in grado di riconoscere un volto necessitava dell'intervento di un operatore che indicasse sulla foto la posizione delle caratteristiche principali del viso di una persona (come occhi, naso, orecchie e bocca). Dopodiché il sistema provvedeva a calcolare distanze e rapporti confrontandoli con dei dati di riferimento.

Il primo vero e proprio sistema automatico di riconoscimento facciale si può datare nel 1988, quando Kirby and Sirovich applicarono il *principal component analysis* (PCA), una tecnica di algebra lineare standard, al problema del riconoscimento. [33]

Negli anni le tecniche di riconoscimento facciale sono migliorate molto, fino ad arrivare al gennaio 2001, quando il primo sistema di prova, che ha poi catturato l'attenzione dei media, è stato installato al Raymond James Stadium di Tampa in occasione di uno dei più grandi eventi americani, il Super Bowl. Il sistema catturava le immagini dai sistemi di sorveglianza e le confrontava con un database di foto segnaletiche digitali creato dal dipartimento di polizia di Tampa, dal Florida Department of Law Enforcement e dall'FBI. In quell'occasione sono stati rilevati 19 piccoli criminali che combaciavano con il database, ma nessuno di essi fu fermato o detenuto in quanto il sistema di riconoscimento facciale era solamente una prova.

L'aggiunta del riconoscimento facciale al sistema di video sorveglianza solleva numerosi dibattiti sul problema dell'equilibrio tra privacy e sicurezza pubblica. Attentati terroristici come quello al World Trade Center, però, hanno alimentato il pensiero comune del bisogno di un sistema che aiuti a difendere la gente dalla libera circolazione di potenziali criminali. Oggi sistemi di riconoscimento

automatico del volto vengono impiegati per identificare bambini scomparsi, minimizzare il furto d'identità, ma soprattutto monitorare il flusso di persone negli aeroporti. [13]

Per rendere gli aeroporti più funzionali e sicuri l'Organizzazione Internazionale dell'Aviazione Civile (ICAO), un'agenzia autonoma delle Nazioni Unite (ONU), che tra le altre cose è incaricata di promuovere la progettazione e lo sviluppo del trasporto aereo internazionale rendendolo più sicuro e ordinato [38], ha stabilito che tutti i passaporti rilasciati dopo il 24 novembre 2005, che non siano a lettura ottica (MRTD), dovranno avere come data di scadenza una data che non superi il 24 novembre 2015. Ciò significa che tali passaporti saranno fuori circolazione entro tale data. [28]

I passaporti a lettura ottica contengono dei chip che hanno memorizzato al loro interno tutti i dati anagrafici e biometrici del possessore in forma digitale. Grazie a questi chip delle apposite macchine possono leggere tutti i dati automaticamente senza l'ausilio di un operatore. Queste macchine conosciute come e-Passport Gates (o *Automatic Border Control*) vengono installate negli aeroporti e permettono di velocizzare i controlli (figura 0.1).



Figura 0.1: Esempio di e-Passport Gates

Una volta acquisita l'immagine del volto all'interno del passaporto viene scattata una foto al possessore del passaporto. Oltre a verificare che il soggetto non sia presente nel database dei ricercati, vengono confrontate le due foto acquisite per verificare che il soggetto possessore del passaporto sia effettivamente la stessa persona rappresentata nella foto del passaporto.

Questi sistemi di riconoscimento facciale possono avere delle difficoltà quando in gioco entrano fattori come l'invecchiamento del soggetto rispetto alla foto del passaporto, un'eventuale chirurgia plastica del soggetto o l'inserimento nel passaporto di una immagine morphed creata appositamente per superare i controlli.

Ma una tecnica molto semplice da mettere in atto per confondere tali sistemi è l'uso di makeup. Attraverso l'uso di cosmetici una persona può evidenziare alcune caratteristiche del volto e celarne altre. L'uso del makeup ai giorni nostri è ormai diventato qualcosa di socialmente accettabile e di uso diffuso, quindi non è strano vedere persone truccate anche in luoghi come aeroporti.

I tipi di attacchi che si possono verificare a tali sistemi tramite l'utilizzo di makeup sono fondamentalmente due: *spoofing* e *obfuscation* (Figura 0.2).

Con *spoofing* s'intende il tentativo di un soggetto di assomigliare a qualcun altro tramite l'uso di makeup. L'obiettivo di questa tecnica è quello di entrare in possesso di un passaporto di una persona non ricercata e di truccarsi in modo tale da sembrare la stessa persona, ingannando così il sistema di riconoscimento facciale.

Con il termine *obfuscation* invece si vuole intendere il tentativo da parte di un soggetto di creare alterazioni al proprio volto per cercare di eludere il rilevamento da parte di sistemi biometrici.

L'obiettivo di questo lavoro di tesi è verificare se l'uso di make-up per eludere i sistemi biometrici può risultare efficace e in tal caso quali contromisure possono essere adottate.

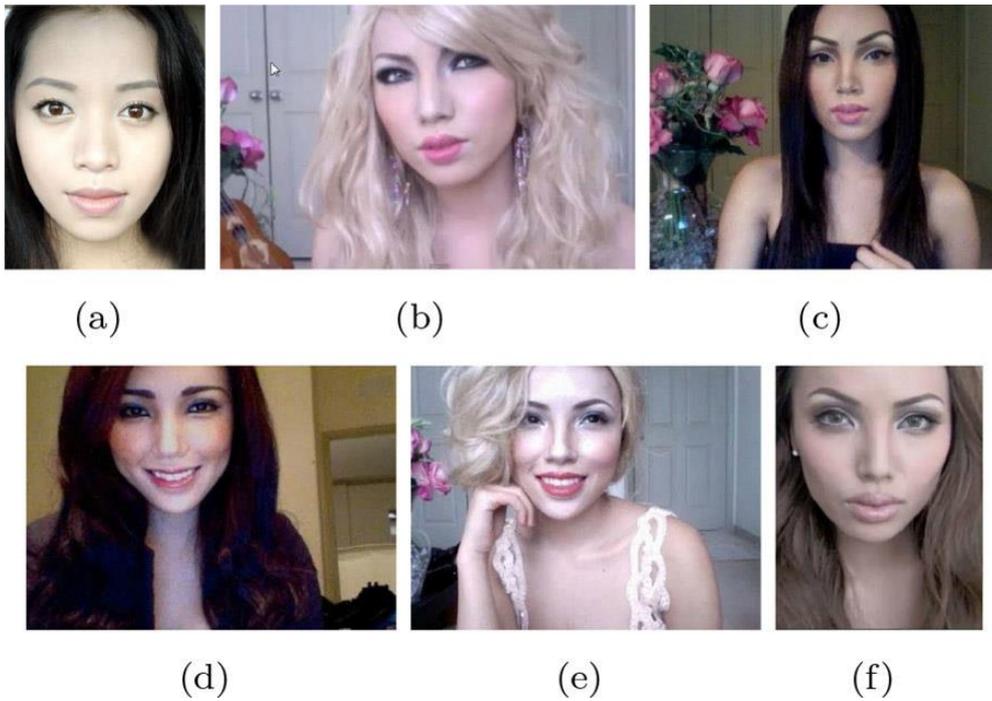


Figura 0.2: Impatto del make-up. (a) soggetto privo di make-up, (b-f) soggetto dopo l'applicazione di make-up (immagini tratte da YouTube)

La tesi è organizzata come segue:

Nel primo capitolo verranno illustrati i passi fondamentali nel processo di riconoscimento del volto, focalizzando l'attenzione sulla localizzazione del volto di una persona in un'immagine, l'estrazione delle caratteristiche principali di un volto e l'uso degli algoritmi di riconoscimento più noti.

Nel secondo capitolo si centra l'attenzione su come i volti possono subire cambiamenti tramite l'applicazione di cosmetici, l'uso di software per inserire in una foto del makeup in maniera digitale e la presentazione dei vari database che verranno poi usati negli esperimenti; si riporta inoltre un primo esperimento per verificare quanto l'applicazione di makeup possa mettere a dura prova un semplice sistema di riconoscimento del volto che non adotti accorgimenti sul tema trattato.

Nel terzo capitolo vengono presentati gli algoritmi messi a punto per far fronte al problema dell'applicazione di makeup nel processo di riconoscimento facciale divisi in categorie (algoritmi che rilevano il trucco, che lo rimuovono e di confronto).

Capitolo 1

Riconoscimento del volto

Un sistema di riconoscimento biometrico è un particolare tipo di sistema informatico che ha l'obiettivo di raccogliere determinate caratteristiche biologiche (impronte digitali, retina, colore e dimensione dell'iride e fisionomia del volto) o comportamentali (impronta vocale, firma, movimenti del corpo) di una persona e confrontarle attraverso degli algoritmi con dei dati precedentemente collezionati in un database. [39]

Un sistema biometrico basato sul confronto delle impronte digitali è un sistema robusto ai cambiamenti e molto attendibile, ma richiede che la persona da riconoscere sia disponibile o costretta a rilasciare le proprie impronte. Al contrario un sistema basato sul riconoscimento del volto è molto sensibile a numerosi fattori quali età, chirurgia plastica, trucco cosmetico e occhiali, ma al contempo molto pratico. Con lo sviluppo di tecniche e tecnologie, oggi un sistema biometrico basato sul riconoscimento del volto riesce ad estrapolare da una sequenza video o da immagini il volto delle persone e tramite avanzati algoritmi li confronta con un database già in possesso (come ad esempio foto segnaletiche di criminali).

I sistemi di riconoscimento biometrico basati sul volto possono avere diverse applicazioni. Oltre al già citato impiego nella sicurezza di luoghi pubblici, sistemi di riconoscimento facciale possono far parte anche della nostra vita quotidiana. Ne sono un esempio il sistema di tag di Facebook, il Kinect di Microsoft sviluppato per il gaming, fotocamere, robotica e in un futuro potranno vedere l'impiego nella sicurezza dei nostri dispositivi permettendo l'accesso esclusivamente dopo aver riconosciuto il volto del proprietario.

Il procedimento di riconoscimento del volto può essere diviso in diverse fasi:

- **Pre-elaborazione:** ci si assicura che l'immagine a cui verrà applicato l'algoritmo di riconoscimento facciale soddisfi dei requisiti standard, come ad esempio la posizione spaziale del soggetto nell'immagine, che la luminosità durante l'acquisizione dell'immagine sia ottimale etc. Questa fase può essere, nella maggior parte dei casi, gestita in maniera automatica da delle apparecchiature utilizzate per il rilevamento delle immagini.
- **Individuazione del volto:** in questa fase il sistema rileva se nell'immagine sottoposta ad elaborazione vi è presente il volto di una persona.
Per prima cosa si individuano, attraverso algoritmi che rilevano il colore della pelle, zone che potrebbero contenere un volto, poi si classificano queste aree come aree contenenti un volto o meno.
La fase di individuazione di un volto può risultare un processo relativamente semplice quando l'immagine viene acquisita in un ambiente controllato (luci, ombre, distanze, contrasti etc.), ma può diventare molto complessa in casi come il rilevamento di volti attraverso un video. In questo caso non è detto che il soggetto da rilevare sia in una posa favorevole allo scopo.
- **Ottimizzazione dell'immagine acquisita:** questa fase è molto importante in quanto semplifica di molto il lavoro delle fasi successive, aumenta le possibilità di successo e diminuisce i tempi di elaborazione. Tramite una

fase di ottimizzazione ci si assicura che l'algoritmo di riconoscimento di un volto verrà applicato ad un'immagine che soddisfi i requisiti richiesti da quel particolare algoritmo (immagine ritagliata attorno al solo volto, soggetto al centro dell'immagine etc.).

- Estrazione delle feature: l'estrazione delle feature (caratteristiche salienti del volto) consiste nell'identificare in un volto individuato in un'immagine, dei particolari specifici di quel volto. Questo passaggio è fondamentale in quanto si vanno a cercare quelle caratteristiche fondamentali che rendono un volto unico e distinto dagli altri (come ad esempio sopracciglia, occhi, naso e labbra).

Una volta individuate, le caratteristiche di un volto vengono rappresentate in forma matematica, così da poter essere processate e confrontate velocemente con quelle di altri volti per stabilire se i due volti in esame appartengono alla stessa persona o a due persone differenti.

- Riconoscimento: la fase di riconoscimento prevede il confronto tra la rappresentazione matematica delle *feature* di un volto con quelle dei volti presenti in un database. Per effettuare tali confronti vengono utilizzati diversi algoritmi.

Di seguito vedremo queste fasi in dettaglio analizzando le tecniche più diffuse.

1.1 Individuazione del volto

Nel caso in cui l'immagine in input non contenga solamente un volto, ma delle persone e uno sfondo, è necessaria una fase di rilevamento del volto. In questa specifica fase viene analizzata l'intera immagine e viene rilevata la presenza e la posizione di eventuali volti all'interno di essa (figura 1.1).

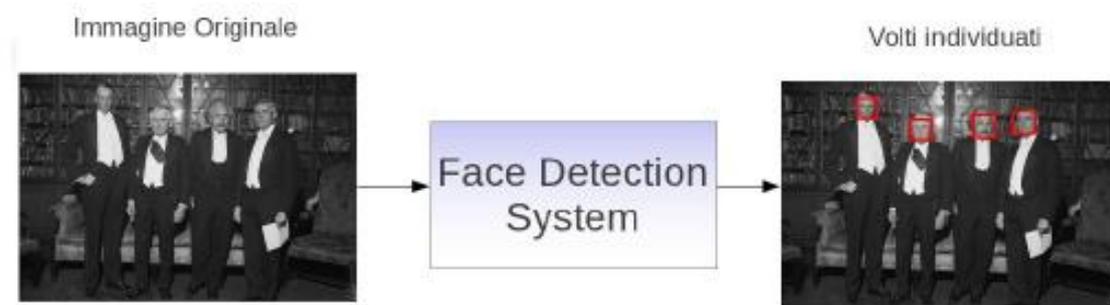


Figura 1.1: Immagine prima e dopo l'individuazione del volto

L'analisi di una intera immagine però potrebbe richiedere troppo tempo. Una tecnica che può ridurre in maniera notevole questo problema è quella di restringere la ricerca di volti a delle sotto-aree che potrebbero contenere un viso. Questa operazione si può implementare cercando delle aree contenenti dei pixel riconducibili al colore della pelle umana. Una volta individuate queste sotto-aree vi si applica poi l'algoritmo per l'individuazione del volto vera e propria. Proponiamo ora le tecniche più conosciute per raggiungere questo obiettivo.

1.1.1 Skin Segmentation

Questa tecnica si basa sul particolare colore che ha la pelle umana. Partendo dagli studi di Brand e Mason [6], Igawa [26] scoprì che la pelle umana è caratterizzata da un particolare rapporto tra il livello di rosso e di verde. In particolare il livello di rosso è predominante e dividendo il valore del livello di rosso per quello di verde (normalizzazione) si riesce ad ottenere un'indicazione approssimativa della presenza di pelle umana nell'immagine.

Per migliorare il processo di individuazione di pelle umana i pixel possono essere filtrati scartando tutti quelli che non rientrano nel range:

$$L_i < R/G < L_s$$

Conducendo numerosi esperimenti usando questa specifica tecnica si è scoperto che il rapporto R/G di pixel rappresentanti pelle umana hanno un limite inferiore pari a 1.

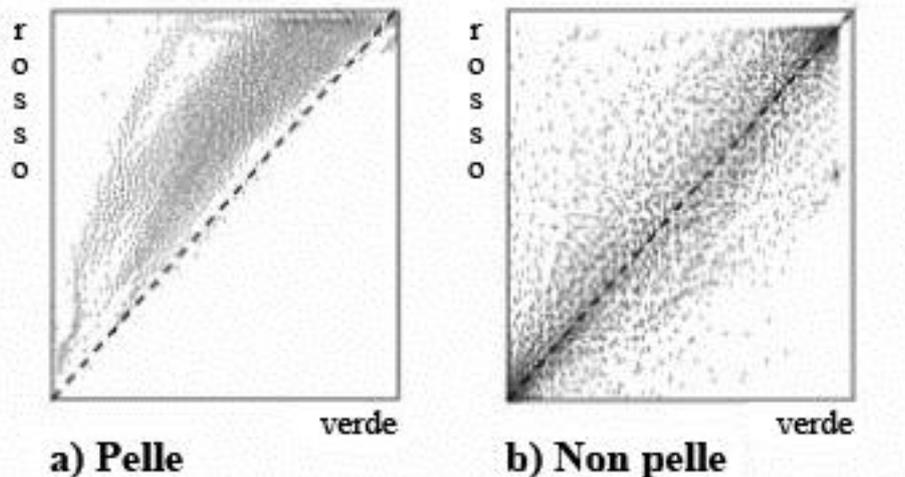


Figura 1.2: Grafici dei pixel di una zona di pelle (a) e non pelle (b)

La figura 1.2 mostra due grafici che rappresentano pixel di una zona contenente pelle umana (figura 1.2-a) e una che non contiene pelle (figura 1.2-b). In questi grafici ogni punto rappresenta un pixel, mentre ascisse e ordinate rappresentano rispettivamente il livello di verde e rosso contenuti in quel particolare pixel. Come processo di scrematura iniziale di zone non contenenti pelle, questo è un processo molto efficiente anche se come si può vedere dal grafico di figura 1.2-b con questa tecnica vengono accettati numerosi pixel non pelle.

1.1.2 Metodo Viola-Jones

Il metodo proposto nel 2001 da Paul Viola e Michael Jones (chiamato metodo Viola-Jones [41]), introdotto inizialmente per la localizzazione generica di

oggetti, è stato applicato con successo nella localizzazione dei volti. Il metodo prevede di creare un classificatore che viene inizialmente addestrato tramite multiple istanze della classe da individuare (esempi positivi) e con immagini che non contengono l'oggetto che si sta cercando (esempi negativi).

L'algoritmo alla base del metodo Viola-Jones classifica delle sotto-finestre dell'immagine in input utilizzando delle *feature* chiamate *Haar-like* (chiamate così per la loro somiglianza con le *Haar wavelet*). Queste *feature* sono costituite da due o più aree rettangolari composte da pixel dalle quali vengono estratte delle informazioni numeriche. Queste aree di pixel possono essere bianche o nere e l'informazione numerica che viene estratta è data dalla differenza tra la somma dell'intensità di pixel dell'area nera e la somma dell'intensità dei pixel nell'area bianca (figura 1.3).

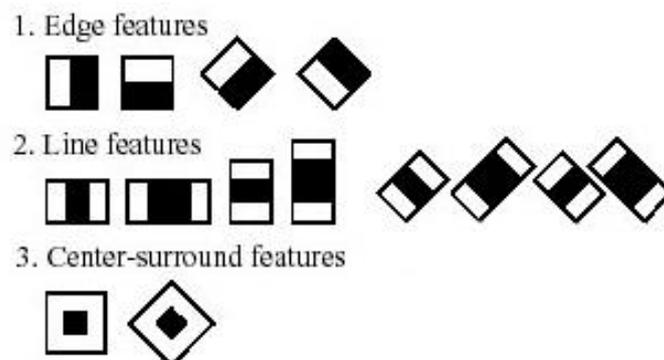


Figura 1.3: Esempi di feature Haar-like

Con questo procedimento le *feature* codificano il contrasto locale nell'immagine. Ricalcolare tutte le volte da zero le *feature* richiederebbe molto tempo, così per ottimizzare l'operazione si fa uso di un'immagine integrale che ha le stesse dimensioni dell'originale, ma i cui valori dell'intensità sono calcolati come la somma di tutti i pixel che si stanno considerando con tutti i pixel sopra e alla sinistra di essi (figura 1.4). Usando questa strategia, il calcolo viene effettuato in

un tempo costante, facendo quattro accessi alla memoria e tre operazioni di somma.

Ad esempio, in riferimento alla figura 1.4, per calcolare la somma all'interno del rettangolo D basta eseguire la seguente formula:

$$I(4) + I(1) - I(2) - I(3)$$

dove I sta per immagine integrale.

Una volta estratte, le *feature* vengono passate ad un classificatore che ha il compito di stabilire a quale classe appartiene il vettore che gli è stato passato: volto o non-volto.

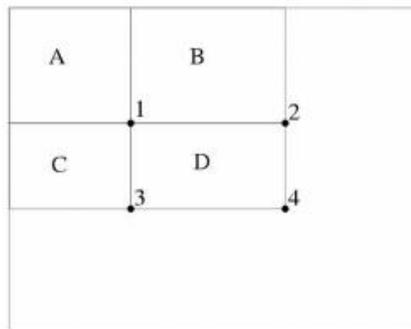


Figura 1.4: Calcolo dell'intensità dei pixel nell'immagine integrale

L'addestramento del classificatore può richiedere anche dei giorni, ma al contrario la localizzazione avviene in tempo reale. L'algoritmo usato per estrarre le *feature* e addestrare il classificatore è una variante di Adaboost.

1.2 Ottimizzazione

L'elaborazione di un'immagine ed il riconoscimento dei suoi pattern sono processi altamente sensibili alla qualità dell'immagine in input, per questo

motivo la fase di ottimizzazione riveste un ruolo essenziale per ottenere dei risultati il più attendibili possibile.

Nel riconoscimento facciale sono molti i fattori che possono influenzare l'analisi di un volto. L'illuminazione è uno dei fattori più compromettenti per l'analisi di un volto, infatti l'eccessiva illuminazione o l'elevata presenza di zone d'ombra possono creare degli effetti ottici che falsano la rilevazione delle caratteristiche da estrarre da un volto. Anche la presenza di accessori quali occhiali o cappelli che coprono parzialmente la faccia causano problemi al sistema di riconoscimento. Un altro problema importante è quello dell'orientazione di un volto.

In questa tesi, per ovviare a tutti questi problemi, verranno usati dei database nei quali le immagini sono state scattate in degli ambienti il più controllati possibile e dove le immagini sono pre-elaborate in maniera tale da doversi preoccupare solamente di applicarvi gli algoritmi eliminando a priori gli inconvenienti descritti.

La prima cosa che può essere fatta per migliorare le prestazioni del sistema è quella di ritagliare e ridimensionare la foto, così facendo il tempo di elaborazione diminuisce.

Subito dopo si passa ad individuare i *landmark* (figura 1.5), ovvero quei punti contraddistintivi di un volto che aiutano a delineare la forma del volto stesso e di tutte le sue caratteristiche principali come naso, occhi e bocca.



Figura 1.5: Immagine dopo l'applicazione dei landmark

Il passo successivo prevede il raddrizzamento del volto, in quanto un volto non bene allineato può causare delle difficoltà nel rilevamento delle aree che caratterizzano le *feature*. Siccome questa fase viene dopo il rilevamento dei *landmark*, conoscendo la posizione dell'iride dell'occhio destro e sinistro, è facile raddrizzare l'immagine allineando gli occhi (figura 1.6).

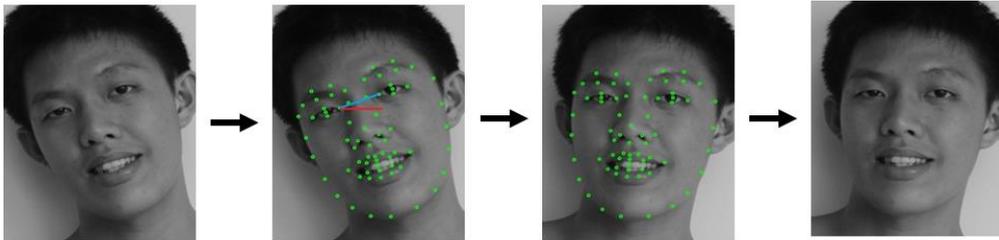


Figura 1.6: Allineamento degli occhi

Nel caso l'immagine sia stata acquisita in un ambiente non controllato, potrebbe essere necessario agire per migliorare il colore. Questo perché ci si potrebbe trovare a lavorare con la qualità dei colori non consona all'analisi del volto, come ad esempio il rilevamento dei contorni delle caratteristiche principali di un volto. A questo proposito ad esempio si può procedere con l'equalizzazione dell'immagine aumentandone così il contrasto (figura 1.7).



Figura 1.7: Aumento del contrasto in una foto

1.3 Da immagine a vettore

Quando si vogliono analizzare una grande quantità di dati sui quali è possibile vi sia dispersione e ridondanza, risulta essenziale estrarre solamente i dati di interesse ed effettuare i calcoli solamente su di essi. Questa operazione viene chiamata estrazione delle *feature* e serve per estrarre da un'immagine le informazioni utili (come le informazioni che derivano dalla codifica di occhi, bocca e lineamenti del volto) in maniera ridotta e codificata in formato numerico. Per ottenere queste informazioni numeriche possono essere usate tecniche di estrazione delle *feature* e/o tecniche di riduzione di dimensionalità. Si possono anche applicare diverse tecniche combinandole assieme per una maggiore efficienza.

Volto visto come un vettore

L'immagine di un volto per essere meglio elaborata può essere trasformata in un vettore (figura 1.8). Se indichiamo con w la larghezza in pixel dell'immagine e con h la sua altezza, il numero delle componenti che si ottiene nella trasformazione è dato dal prodotto $w*h$, dove ogni pixel dell'immagine iniziale corrisponde ad una componente del vettore. Per costruire tale vettore è sufficiente quindi concatenare le righe della matrice dell'immagine iniziale.



Figura 1.8: Costruzione di un vettore

Ovviamente in questo esempio nel vettore vengono memorizzate informazioni ridondanti e non necessarie al riconoscimento di un volto. Per evitare questo bisogna effettuare prima un processo di estrazione delle *feature*.

Di seguito verranno presentate le tecniche più utilizzate di estrazione delle *feature*, di riduzione di dimensionalità e i classificatori più noti.

1.3.1 Estrazione delle feature

Di seguito verranno analizzate le tecniche più comuni di estrazione di *feature* da un volto.

Local Binary Pattern

Il Local Binary Pattern è un tipo di *feature* molto utilizzato nel riconoscimento facciale partendo da immagini 2D *grayscale* per via della sua semplicità computazionale e la sua robustezza rispetto a variazioni di illuminazione.

Nella sua versione base ad ogni pixel viene associato un valore binario dipendente dalla valorizzazione dei pixel che stanno nell'intorno 3x3 di quello preso in esame: se il valore del pixel centrale è maggiore di quello del pixel dell'intorno che si sta esaminando, allora al pixel dell'intorno si attribuisce il valore 1, 0 altrimenti. Mettendo in sequenza (in senso orario) i valori binari ottenuti confrontando il pixel centrale con i pixel attorno ad esso (partendo dal pixel in alto a sinistra), si ottiene un numero binario di 8 cifre, che convertito poi in decimale rappresenta il valore finale associato al pixel di riferimento (figura 1.9).

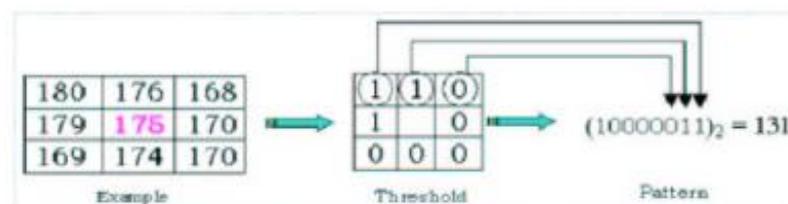


Figura 1.9: Calcolo dell'operatore LBP

Nella sua versione generalizzata di LBP, ogni pixel viene considerato in un intorno circolare di p punti con raggio r (figura 1.10). Se gli elementi non ricadono su un unico pixel, i loro valori sono ottenuti per interpolazione. Generalmente il numero di pixel vicini scelto è 8. Considerare intorni circolari e utilizzare l'interpolazione per dare un preciso valore a ciascun punto campione dell'intorno non cambia il processo di costruzione del valore binario associato al pixel centrale. Il procedimento utilizzato per l'intorno 3x3 vale anche con l'intorno circolare.

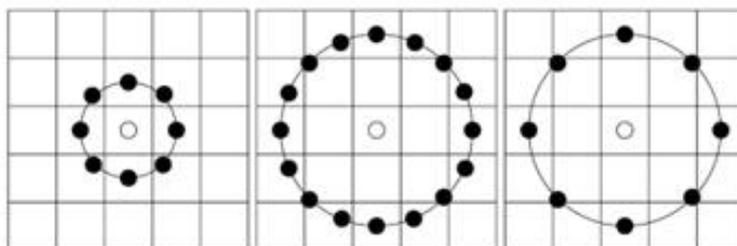


Figura 1.10: Calcolo LBP con diversi vicini e pixel. (8,1) (16,2) e (8,2)

Con i numeri ottenuti si calcolano gli istogrammi delle frequenze, eventualmente si normalizzano e poi si concatenano gli istogrammi di tutte le celle ottenendo così un vettore delle *feature*.

Filtri di Gabor

I filtri di Gabor vengono spesso utilizzati per l'estrazione di caratteristiche di un volto grazie alla loro proprietà di localizzazione sia nel dominio del tempo che della frequenza. Sono filtri lineari in grado di descrivere la distribuzione spaziale dei livelli di grigio di una determinata regione. Un filtro di Gabor (figura 1.11-c) si ottiene attraverso la modulazione di una funzione Sinusoidale (figura 1.11-a) con una Gaussiana di deviazione standard (figura 1.11-b).

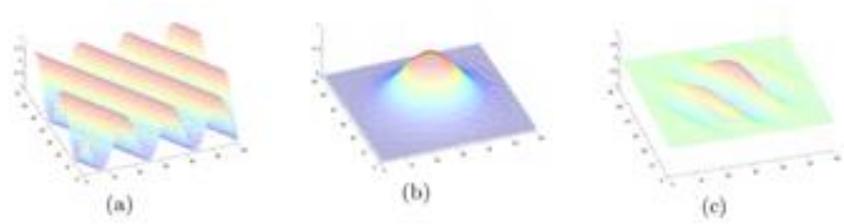


Figura 1.11: (a) funzione sinusoidale, (b) Gaussiana e (c) filtro di Gabor

Il filtro di Gabor è dato dalla seguente equazione:

$$g(x, y, \theta, \phi) = \exp\left(-\frac{x^2 + y^2}{\sigma^2}\right) \exp(i(2\pi\theta(x\cos\phi + y\sin\phi)))$$

Dove x e y indicano le coordinate spaziali nell'immagine, mentre θ e ϕ indicano scala e orientamento.

Il risultato dell'equazione è un filtro di Gabor visto come due immagini che identificano parte reale e immaginaria del filtro.

1.3.2 Dimensionality Reduction

Lo scopo delle tecniche di riduzione di dimensionalità è quello di trasformare i dati da uno spazio iniziale, ad uno di dimensioni inferiori. I metodi più usati sono PCA (Principal Component Analysis) [42], LDA (Linear Discriminant Analysis) [36] e CCA (Canonical Correlation Analysis [43]). Sono metodi lineari di analisi statistica che permettono di trovare dei pattern in un insieme di grandi dimensioni classificandoli in base al loro grado di similarità.

PCA

Principal Component Analysis esegue un mapping lineare cercando di ridurre al massimo le dimensioni dello spazio di partenza mantenendo allo stesso tempo l'informazione originale. Attraverso la combinazione lineare delle variabili originali, si ottengono delle nuove variabili che possano spiegare una porzione della varianza totale dei dati. Vengono poi calcolati gli autovettori della matrice di covarianza e solo quelli corrispondenti ai valori più alti di varianza vengono conservati, mantenendo limitata così l'analisi alle principali variabili.

LDA

LDA, conosciuto anche con il nome di Fisher Discriminant Analysis, è un altro metodo di riduzione della dimensionalità dello spazio. Mentre PCA privilegia le dimensioni che codificano e rappresentano i pattern, LDA privilegia quelle che al meglio le discriminano.

CCA

CCA permette, avendo due vettori, di trovare delle correlazioni tra le variabili producendo in tal caso una combinazione lineare con la massima correlazione.

1.4 Riconoscimento

Una volta ottenute le *feature* che descrivono un volto, si passa alla fase di riconoscimento durante la quale le *feature* provenienti da due immagini differenti vengono messe a confronto per stabilire se le *feature* estratte provengano dalla stessa persona.

Oltre ad algoritmi di riconoscimento, questa fase si avvale di una fase di addestramento. In questa fase si utilizzano un insieme di dati di training per far sì che il sistema impari a raggruppare in determinate categorie i dati in base alle loro caratteristiche principali (*feature*).

1.4.1 Support Vector Machine

Support Vector Machine (SVM) viene utilizzato, nei sistemi di riconoscimento facciale, come classificatore. E' composto da un insieme di metodi di apprendimento supervisionato per la regressione e la classificazione di pattern. E' un classificatore lineare generalizzato, quindi se le classi sottoposte ad analisi non sono linearmente separabili, viene eseguito un *mapping* dei pattern su di uno spazio di dimensione superiore, così da poter separare le classi avendo maggiori gradi di libertà.

SVM è ideato come classificatore binario, permettendo quindi di raggruppare i pattern in due grandi classi. Esistono comunque approcci per risolvere il problema della classificazione con più di due classi.

1.4.2 Eigenfaces

Eigenfaces, conosciuto come il metodo delle autofacce, è un metodo di estrazione delle *feature* basato su PCA. Dall'estrazione di queste *feature* si ricavano degli autovettori chiamati appunto *eigenfaces* che vengono usati poi come base per il nuovo spazio.

Ogni *eigenface* creata descrive caratteristiche diverse di un volto: la linea dei capelli, la simmetria, la larghezza del naso (figura 1.12). Grazie a questa loro capacità di esprimere caratteristiche diverse di un volto, la faccia di un individuo

può essere rappresentata come la combinazione lineare delle *eigenfaces*, oppure essere approssimata utilizzando solo le *eigenfaces* più significative.

Durante la fase di addestramento, l'algoritmo ha complessità pari a $O(n^2m^2L)$, mentre nella fase di test si riduce a $O(LMN)$. Dove n e m rappresentano il numero di colonne e righe dell'immagine considerata come matrice, mentre L è il numero di vettori proiezione, N il numero di elementi della fase di addestramento e M degli elementi nella fase di test.

Il metodo descritto è ottimo in un ambiente controllato nel quale si possono gestire posa, illuminazione ed altri fattori che possono ridurre l'efficienza durante un processo di riconoscimento del volto. In caso i volti non siano catturati in posa frontale o nelle stesse condizioni (basta pensare a semplici casi come diversa illuminazione o applicazione in una delle immagini di makeup), il sistema perde la sua efficacia rendendo necessario indirizzarsi su altre tecniche più robuste.



Figura 1.12: Esempio di volti utilizzati durante una fase di addestramento (a) e le relative eigenfaces (b)

1.4.3 Hidden Markov Model

I modelli di Markov sono uno strumento statistico in grado di riprodurre l'evoluzione di un sistema non deterministico mediante l'uso di stati. Le

probabilità associate ad ogni transizione sono memorizzate in una matrice. I valori vengono calcolati attraverso l'osservazione dello storico del sistema. Assieme alla matrice descritta è necessario l'uso di un vettore che descriva le probabilità iniziali.

HMM risultano fondamentali quando alcuni stati non siano direttamente osservabili (da cui il nome hidden state). Siccome i tratti somatici del volto di una persona seguono sempre lo stesso schema il viso può essere descritto perfettamente da un modello bidimensionale HMM.

In particolare, secondo HMM, in un volto si possono identificare cinque grandi stati principali: fronte, occhi, naso, bocca e mento (figura 1.13). Ognuno di questi stati può essere poi suddiviso in stati più piccoli che esprimono delle caratteristiche somatiche della persona.

HMM è un modello perfetto per riconoscere un qualunque oggetto che mantenga le proprie caratteristiche in termini di posizione, dall'alto al basso e da sinistra a destra. Con queste premesse risulta intuitivo comprendere quanto HMM possa essere sfruttato nella fase di riconoscimento di un volto.

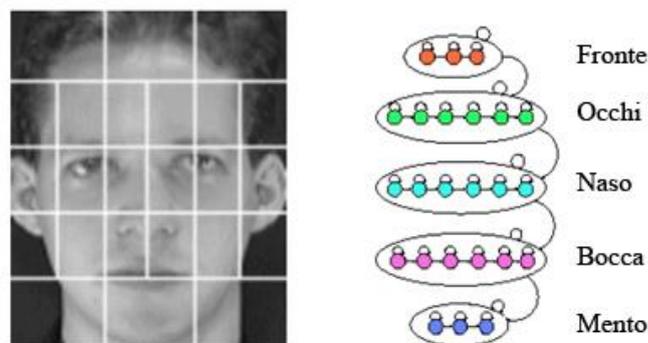


Figura 1.13: Stati del modello HMM

Si parte con una fase di addestramento nella quale vengono archiviati modelli per ogni individuo che si vuole riconoscere. Nella seconda fase si crea un modello dell'individuo che si vuole riconoscere e lo si confronta con quelli memorizzati precedentemente per trovare l'accoppiata con maggiore somiglianza. Questa somiglianza viene espressa in termini probabilistici.

E' stimato che i sistemi che utilizzano HMM abbiano una probabilità di commettere errori di circa il 2-3%.

1.4.4 Adaboost

Adaboost è un classificatore non lineare complesso ottenuto come combinazione lineare di altri classificatori deboli. L'idea di Adaboost è quella di creare una lista di classificatori, assegnando, in maniera iterativa, un peso ad ogni nuovo classificatore che identifichi la capacità di riconoscere in maniera corretta i campioni non riconosciuti invece da altri classificatori già valutati. Infine questi classificatori daranno un voto che avrà incidenza maggiore in base al peso a loro assegnato. La scelta finale avverrà per maggioranza.

$$F_T(x) = f_1(x) + f_2(x) + \dots + f_T(x) = \sum_{t=1}^T f_t(x)$$

dove f_1, \dots, f_T sono i singoli classificatori.

Se il classificatore addestrato non trova un oggetto che è presente, oppure ne identifica la presenza dove invece non c'è, lo si può ricalibrare aggiungendo gli esempi positivi o negativi al suo addestramento.

La *feature* scelta per un dato classificatore è quella che permette di ottenere l'errore di classificazione complessivamente più basso.

Capitolo 2

Makeup

Il makeup è diventato un accessorio di uso comune soprattutto nel mondo femminile. Questo processo di diffusione dell'uso di cosmetici come mezzo di abbellimento di un volto non appartiene alla modernità, infatti vi si possono trovare tracce già nell'Antico Egitto [3] e nell'Antica Grecia [1].

Numerosi studi hanno analizzato come l'utilizzo di cosmetici aumenti considerevolmente l'attrattività di una persona [9] e l'effetto che il makeup applicato sul volto di una donna ha su di un uomo [14]. Questo spiega il motivo per il quale nella società moderna, sempre più fondata su di un modello di bellezza rigido nel quale bisogna assomigliare il più possibile a modelli prestabiliti, vi sia un aumento spropositato nella vendita e nell'uso di cosmetici. Nel 2011 negli Stati Uniti sono stati venduti cosmetici per il makeup (rossetti, ombretti, fondotinta, eyeliner, etc.) per un ammontare di 3,6 miliardi di dollari (3,3 miliardi di euro), aumentando così del 9% rispetto al 2010 [25].

L'applicazione di prodotti cosmetici abbelliscono un volto andando a modificare elementi fondamentali di una persona quali colore della pelle e delle labbra, contorni, contrasti ed eliminando le imperfezioni. Tramite l'uso abbondante di trucco abbinato ad alcune tecniche di applicazione, può risultare difficile per un

essere umano riconoscere ad una prima impressione una persona che si è abituati a vedere struccata. Come si può vedere dalla figura 2.1 l'uso di semplice makeup può facilmente alterare la percezione umana di un volto confondendolo con quello di un altro (nel caso dell'immagine quello di un personaggio famoso) [2].



Figura 2.1: Impatto dell'applicazione di prodotti cosmetici.

Sulla base di questa considerazione si può facilmente intuire che sistemi biometrici basati su delle caratteristiche fisiche, che il trucco va ad alterare, potrebbero essere ingannati.

In questo capitolo analizzeremo quali prodotti possono essere applicati e in quali punti del volto essi potrebbero creare più problemi ad un sistema di riconoscimento facciale.

Verrà analizzato anche l'uso di software in grado di inserire in maniera digitale il trucco su delle immagini di volti. Siccome lo studio degli effetti del trucco su di un sistema di riconoscimento facciale è un argomento abbastanza recente e quindi non si hanno a disposizione molti campioni di volti su cui condurre esperimenti, questi software sono molto importanti in quanto simulano molto bene l'applicazione di un trucco vero e proprio ed in questo modo si possono creare dei database che contengano volti senza e volti con il makeup.

Verrà poi descritta la struttura di ogni database usato per testare i vari algoritmi utilizzati per contrastare l'effetto del makeup sul riconoscimento.

Ed infine un'analisi dettagliata sull'impatto che il trucco ha su di un sistema biometrico di riconoscimento del volto che non implementa delle tecniche ad hoc contro gli effetti del makeup.

2.1 Applicazione del makeup

I volti umani condividono le stesse caratteristiche (occhi, sopracciglia, naso e labbra), ma dimensioni, forma, simmetria [31] e distanze tra di esse giocano un ruolo importante nella percezione della bellezza. Il makeup è uno strumento potente e non permanente che permette di coprire i propri difetti, sembrare più giovani e più attraenti creando un effetto di pelle morbida ed occhi e labbra più grandi [9].

Seguendo la classificazione pensata da Eckert, Kose e Dugelay [11], l'applicazione di makeup può essere diviso in otto passaggi principali. Questi passaggi sono stati pensati osservando i vari tutorial di makeup che si trovano su Youtube e i vari tipi di prodotti che si possono utilizzare (figura 2.2).



Figura 2.2: Zone di applicazione del makeup

Step	Prodotti	Effetti
Base	Fondotinta, Correttore, Cipria	Copertura di difetti della pelle e dei naturali effetti d'ombra
Guancia	Blush, Bronzer	Enfatizza e cambia la percezione della posizione degli zigomi. Dona un effetto più vivace e giovanile.
Sopracciglia	Matita	Accentua, definisce e cambia la percezione della posizione del sopracciglio. Incrementa il contrasto.
Ombre (palpebra superiore)	Ombretto	Enfatizza la palpebra, incrementa la profondità e cambia la percezione della dimensione e del colore degli occhi.
Palpebra superiore	Eyeliner, Kajal (matita per interno occhi)	Valorizza i contorni, cambia la percezione della dimensione degli occhi e ispessisce le ciglia.
Mascara	Mascara, Ciglia finte	Aumenta la densità, curva e allunga le ciglia, crea l'illusione di occhi più grandi.
Palpebra inferiore	Eyeliner, Kajal (matita per interno occhi)	Accentua i contorni e cambia la percezione della dimensione degli occhi
Labbra	Rossetto, Matita, Gloss	Accentua i contorni, cambia la percezione delle dimensioni, del colore e del contrasto delle labbra, dona lucentezza.

Tabella 1: 8 step di applicazione del makeup

La tabella 1 riassume gli otto passaggi principali, i prodotti che si possono usare in ogni passaggio e l'effetto che questi provocano sul volto.

Questi 8 step formano un codice (ogni step è caratterizzato da un numero che va da 1 a 8). Ognuno di questi passi può essere applicato in maniera leggera o pesante e ad essi verrà applicato un codice rispettivamente di 1 o 2, mentre per l'assenza di makeup viene usato il codice 0.

Nonostante le parti di un volto che possono essere truccate siano molte, ai fini del riconoscimento facciale si possono raggruppare in macro aree: pelle, occhi e bocca.

Gli step che coinvolgono guance e fondotinta appartengono alla macro area della pelle, le labbra coinvolgono la bocca e il resto degli step sono incentrati sugli occhi. Dividendo la faccia nelle tre aree citate si capisce subito che per studiare il problema dell'incidenza dei cosmetici nel riconoscimento facciale bisogna focalizzare l'attenzione negli occhi, nella presenza di trucco in grado di cambiare sfumature della pelle e nella bocca, senza stare a concentrarsi in ogni piccolo particolare. Ogni macro area viene considerata in maniera indipendente, in maniera tale che si possa avere una visione in merito alla sensibilità al cambiamento che ogni zona ha in risposta all'applicazione dei cosmetici.

Un'ulteriore classificazione la si può fare in base a quanto e in quale modo il trucco può cambiare la percezione delle dimensioni e del colore delle varie caratteristiche del volto di una persona (Tabella 2). Analizzando numerosi tutorial di makeup si possono distinguere tre principali categorie: leggero, intermedio e pesante, più l'aggiunta di una quarta categoria che rappresenta l'assenza totale di makeup.

Unendo i dati della tabella 1 (codificandola con i numeri da 1 a 8) a quelli della tabella 2 si può creare una classificazione del makeup che tiene conto della relazione tra parte del volto e quantità di trucco applicata, mettendo in evidenza l'effetto che la pesantezza del trucco ha in maniera più o meno marcata (figura 2.3).

Categoria	Descrizione
Nessun Effetto	Makeup non applicato, nessun effetto
Effetto Leggero	Piccoli cambiamenti del colore e nella percezione delle forme
Effetto Intermedio	Forte accentuazione dei contorni e delle caratteristiche di un volto
Effetto Pesante	Perdita delle caratteristiche proprie di un volto, effetto artificiale e di mascheramento

Tabella 2: Pesantezza del trucco applicato e relativo effetto

Le colonne di figura 2.3 rappresentano l'applicazione di trucco in una determinata zona del volto. In ogni colonna sono presenti i numeri da 0 a 2 che come detto rappresentano la quantità di trucco applicata alla zona. La riga dove viene inserito il numero identifica l'impatto che l'applicazione del trucco in una determinata quantità ha sul cambiamento del volto.



Figura 2.3: Effetto del trucco in base alla zona e alla quantità dell'applicazione

Vediamo nel dettaglio, studiando la figura 2.3, come l'applicazione del makeup influisce sulla percezione delle tre macro aree (occhi, pelle e bocca).

- Occhi: se ombretto, mascara e trucchi per la palpebra inferiore vengono applicati in maniera pesante, la zona degli occhi viene categorizzata sotto "trucco pesante", ovvero l'occhio perde molte delle sue caratteristiche originali attribuendogli un effetto artificiale. L'applicazione di trucco in maniera pesante sulle sopracciglia o sulla palpebra superiore genera un effetto di percezione del trucco intermedio in quanto il makeup di queste aree risalta quelli che sono i lineamenti propri dell'occhio della persona. Qualsiasi altro trucco applicato in maniera leggera crea un effetto leggero di impatto visivo.
- Pelle: se i trucchi per la pelle e per le guance vengono usati in maniera pesante, l'alterazione che ne segue viene categorizzata come pesante. Il fondotinta uniforme ed appiattisce le forme del viso, i difetti vengono coperti. L'applicazione di makeup sulle guance altera la percezione di dimensione, colore e posizione degli zigomi. Anche in questo caso si ha un effetto di alterazione artificiale. Si può notare che non vi è del makeup che applicato abbia un effetto di alterazione intermedio per quanto riguarda la pelle.
- Labbra: l'applicazione di makeup pesante alle labbra cambia forma e dimensione delle stesse alterando pesantemente il loro aspetto. L'uso di rossetto, matita e gloss in maniera leggera aumenta il contrasto e i contorni della bocca creando un effetto intermedio di alterazione delle caratteristiche. Si nota che non vi è un effetto leggero di alterazione.

La categoria nella quale rientra il makeup intero di un volto è uguale alla categoria più alta delle macro aree in cui abbiamo suddiviso una faccia. Bisogna tuttavia considerare che le macro aree contribuiscono in maniera differente alla definizione della categoria del makeup dell'intero volto. Il trucco della bocca

genera un piccolo impatto se consideriamo la faccia nel suo insieme. Quindi l'effetto intermedio del trucco della bocca incide in maniera leggera se si considera il volto nella sua interezza.

2.2 Software per il makeup virtuale

I software per il makeup virtuale (conosciuti anche come Virtual Makeover Software, software VMO) sono software e tool online che possono essere usati per applicare del makeup. Prendendo la foto (possibilmente frontale, priva di trucco e senza capelli davanti al volto) della faccia di una persona, con l'aiuto del software è possibile applicare cosmetici, stile e colore dei capelli diversi da quelli naturali (figura 2.4). Nonostante il trucco applicato sia virtuale, riesce comunque ad assumere un aspetto realistico. I software più avanzati danno anche la possibilità di inserire accessori quali lenti a contatto e occhiali. Lo scopo principe dei tool che si trovano online è quello di far provare l'effetto dei cosmetici sul proprio volto senza dover andare fisicamente in negozio. In questo modo, comodamente da casa, una persona può testare il trucco e se di proprio gradimento acquistarlo direttamente online e farselo recapitare al proprio domicilio.

Per quanto riguarda il problema del face recognition analizzato in questo lavoro di tesi, questi software sono molto importanti in quanto permettono di avere a disposizione database più grandi. Per testare meglio gli algoritmi alle prese con volti con e senza makeup questi software sono l'ideale, perché permettono di decidere quale parte del volto truccare e creare più versioni di una stessa faccia, ma con pesantezza di trucco differenti, rendendo lo studio degli algoritmi ancora più esaustivo.

Di seguito verranno presentati i due tool online in ordine crescente di precisione e notorietà.



Figura 2.4: Esempio di Makeup virtuale

2.2.1 Mary Kay Virtual Makeover

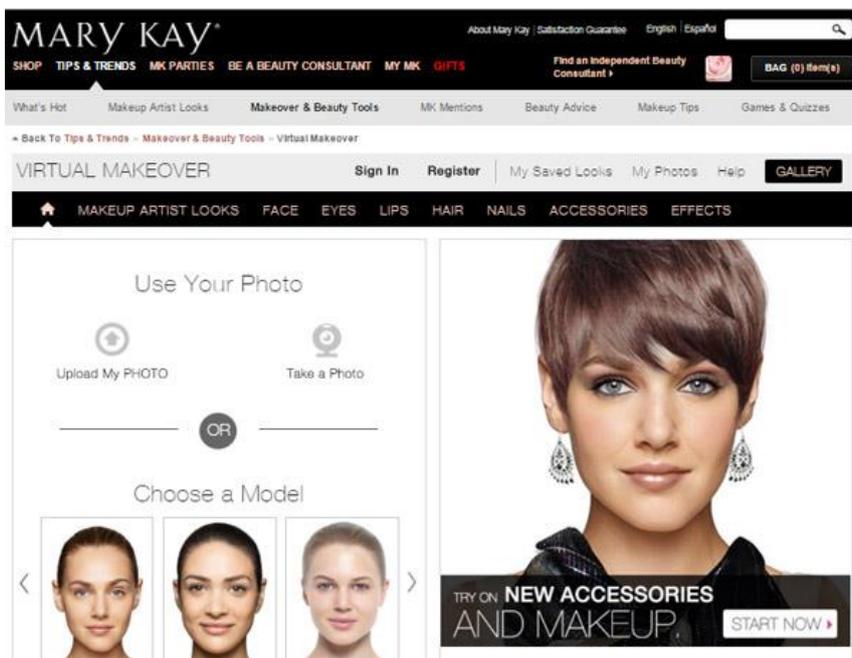


Figura 2.5: VMO Mary Kay [22]

La Mary Kay Inc. è una grande azienda americana che si occupa della vendita diretta di cosmetici e in più di 50 anni ha venduto prodotti in più di 35 paesi nel mondo. Secondo la Direct Selling News, nel 2011 Mary Kay era la sesta più grande compagnia con un fatturato netto di 2,9 miliardi di dollari [40].

Dalla schermata principale del software (figura 2.5) si può selezionare se caricare una foto da computer, scattarsi una foto tramite webcam oppure lavorare su una foto scegliendo tra le modelle messe già a disposizione dal tool. Se si sceglie di caricare una propria foto bisogna assicurarsi che essa sia in formato JPG o PNG e che le sue dimensioni ricadano nel range tra 250 Kb e 4 Mbs. La foto che si vuole caricare deve rispettare possibilmente delle linee guida per permettere al software di identificare il volto nella maniera più semplice possibile: volto privo di trucco (o comunque che non alteri il colore di zone della faccia), nessun accessorio, capelli raccolti e che non coprano il viso, luminosità non troppo scura ma nemmeno troppo alta e infine scattare la foto con uno sfondo ben distinto, preferibilmente una parete bianca. Successivamente si può ritagliare la foto.

Una volta caricata si può scegliere se applicare un look completo già preparato o spostarsi nel menù in alto potendo scegliere di aggiungere trucco al volto, agli occhi, alle labbra e modificare l'acconciatura. Cliccando su ogni categoria si possono scegliere tra sotto-categorie di trucchi (ad esempio nella categoria faccia si può poi scegliere tra fondotinta, correttore e blush).

Se il trucco non viene applicato correttamente, non coprendo adeguatamente le zone, significa che il software non è riuscito ad individuare correttamente tutte le parti del viso. Si può andare manualmente a modificare l'area di ogni zona. Per fare ciò si può cliccare sulla voce *Retrace* posta sotto la foto caricata e scegliere qual parte del volto si vuole ridefinire. In figura 2.6 si può vedere un esempio del trucco applicato alla zona occhi in maniera virtuale.



Figura 2.6: Occhi prima e dopo aver applicato il trucco con il Mary Kay VMO

2.2.2 TaaZ Virtual Makeover

TaaZ è una società molto giovane. Fondata a San Diego nel 2007, TaaZ si è affermata subito nell'ambiente del makeup andando incontro ai bisogni del mondo femminile. La tecnologia usata da TaaZ è stata considerata come una delle più realistiche e facili da usare offrendo un'esperienza di virtual makeover gratuita.

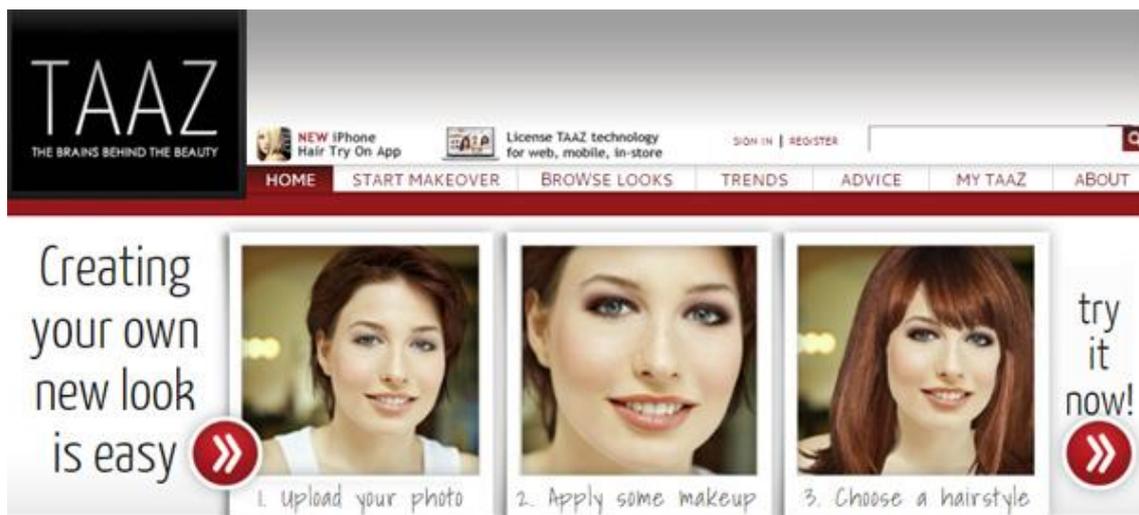


Figura 2.7: Homepage TaaZ.com

Cliccando sulla voce “Creating your own new look is easy” si può iniziare il procedimento di Virtual Makeover (figura 2.7). Come per il software di MaryKay bisogna scegliere se lavorare su di una modella messa a disposizione dal sito o se caricare una foto dal proprio computer. In questo caso non è possibile scattare una foto direttamente da webcam, anzi viene proprio sconsigliato per evitare di usare immagini con condizioni di qualità e luminosità non adatti. I requisiti richiesti dal software sono: immagine almeno 200 x 200 pixel; come formato dell'immagine si ha a disposizione JPG, GIF o PNG; foto frontale di un volto e senza makeup con condizioni di luminosità ottimali, senza zone scure d'ombra.

Una volta caricata la foto vi verrà data la possibilità di dare un nome al vostro lavoro, ruotare l'immagine e ritagliare la vostra foto. Nella schermata successiva dovrete confermare o meno se il programma è riuscito ad individuare correttamente occhi e bocca. Se confermate potrete iniziare ad applicare il makeup, altrimenti partirà una breve configurazione nella quale potrete manualmente indicare forma di occhi, iride, sopracciglia e bocca. Anche in questo caso abbiamo a disposizione le categorie delle zone principali del volto (Faccia, occhi, labbra, capelli, accessori e look completi) e per ognuna di esse si hanno delle sotto-categorie che identificano i vari trucchi applicabili in quella zona. Sono a disposizione molti colori e molti stili di applicazione (per quanto riguarda gli occhi è anche possibile applicare delle lenti o rifinire le sopracciglia). Per qualsiasi zona del volto, una volta applicato il trucco, si può andare a ritoccare i contorni della zona (ad esempio gli occhi) cliccandoci sopra con il mouse per migliorare la precisione nel caso il trucco venga applicato in modo anomalo.

Sopra alla foto è presente un comodo pulsante “BEFORE + AFTER” che permette di vedere a confronto l'immagine caricata senza makeup con il makeup creato fino a quel momento.

Per salvare il lavoro realizzato bisogna prima creare un account su Taaz.com.

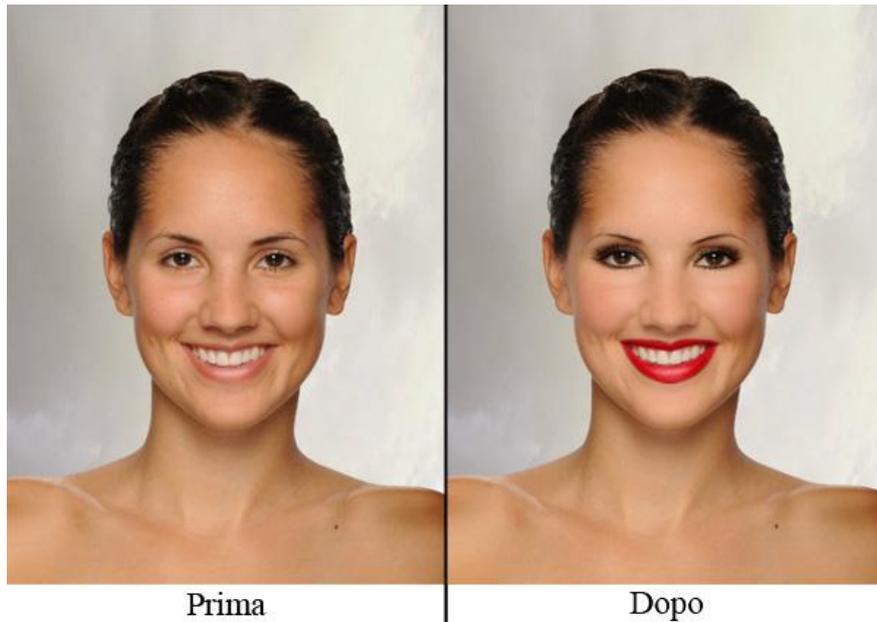


Figura 2.8: Makeup virtuale totale applicato con Taaz VMO

Vista la qualità ed i risultati ottenuti (figura 2.8), il tool online di Taaz.com risulta perfetto per creare dal nulla (avendo solo a disposizione foto di volti senza makeup) un database che per ogni volto ci permette di avere più tipi di makeup, riuscendo così a testare in maniera più precisa gli effetti del trucco su di un sistema di riconoscimento facciale.

Nel prossimo paragrafo analizzeremo tutti i database utilizzati per testare i vari algoritmi che tentano di mitigare gli effetti del trucco e tra questi anche il database creato con Taaz.

2.3 Database

Per database intendiamo delle collezioni di immagini con certe caratteristiche che possano semplificare il lavoro di ricerca mettendo a disposizione dei campioni già pronti per poter testare gli algoritmi. Nei database di seguito descritti vengono collezionate immagini di persone con e senza makeup. Questi database

sono fondamentali in quanto verranno poi utilizzati negli esperimenti necessari per testare i vari algoritmi di riconoscimento facciale in presenza di makeup. Una volta testati, gli algoritmi potranno essere confrontati con una precisione maggiore analizzando i dati provenienti dagli algoritmi che hanno usato gli stessi database.

Di seguito vedremo in dettaglio diversi database:

- Youtube Makeup database;
- Makeup In the Wild database;
- Facial Cosmetics Database;
- Virtual Makeup database;
- FAcE Makeup database.

2.3.1 Youtube Makeup database



Figura 2.9: Campione di immagini tratti dal YMU. Le immagini (a) e (c) rappresentano volti senza makeup, mentre (b) e (d) dopo l'applicazione di makeup.

Youtube Makeup database (YMU) è un database creato estrapolando le immagini dai tutorial di makeup presenti su Youtube. Il database ritrae 151 soggetti femminili di razza caucasica. Per ogni soggetto sono state inserite nel database due immagini in assenza di trucco e due immagini dopo l'applicazione di makeup, collezionando così quattro foto per ogni soggetto (esempi in figura 2.9). Per alcuni soggetti sono state acquisite tre immagini prima e dopo il trucco mostrando così una graduale variazione di trucco, da leggero a pesante. In totale il database contiene 604 immagini, 302 rappresentanti soggetti privi di makeup e 302 dopo l'applicazione al volto di cosmetici.

Le foto dei soggetti, essendo state rilevate da dei video tutorial, non sono state acquisite in un ambiente controllato per cui le espressioni e le pose possono variare da una foto all'altra. L'illuminazione invece la si può considerare costante in quanto le immagini di una persona sono state acquisite durante lo stesso tutorial.

L'alterazione di percezione di un volto causata dal trucco si concentra sulla qualità della pelle e sul colore delle labbra, ma il cambiamento maggiore provocato dall'applicazione di cosmetici è dovuto al trucco nella zona occhi, dove questi ultimi possono essere accentuati tramite l'utilizzo di diversi prodotti.

In alcuni casi la pettinatura cambia drasticamente prima e dopo l'applicazione del makeup.

Questo database è stato creato da Dantcheva [10] per testare il suo lavoro. In quell'occasione sono stati usate le immagini di 99 soggetti, mentre nel suo lavoro successivo [8] utilizzò tutti i soggetti a disposizione.

2.3.2 Virtual Makeup database



Figura 2.10: Campione di immagini tratti dal VMU. (a) volto senza makeup, (b) solo occhi truccati, (c) solo bocca truccata, (d) makeup totale.

Il Virtual Makeup database (VMU) è stato creato partendo da 51 immagini di soggetti femminili di razza caucasica senza trucco, estrapolati dal Face Recognition Gran Challenge database (FRGC). Da ogni immagine ne sono state create tre applicandovi virtualmente del makeup, per un totale di 4 immagini per ogni soggetto.

Questo database è stato creato per mettere alla prova, assieme al YMU, gli studi di Dantcheva [10]. Per aggiungere il trucco ai volti a disposizione è stato utilizzato Taaz, il tool descritto nel paragrafo 2.2.2. Come si vede chiaramente da Figura 2.10 per ogni soggetto si ha a disposizione una foto del soggetto privo di trucco (figura 2.10-a), una in cui al soggetto è stato applicato del makeup solo nella zona occhi (figura 2.10-b), una con solo il makeup alle labbra (figura 2.10-c) e infine una a cui è stato applicato un total makeup (figura 2.10-d).

2.3.3 Makeup In the Wild database



Figura 2.11: Campione di immagini tratte dal MIW. Sopra soggetti senza makeup, sotto soggetti con makeup.

Makeup in the “wild” database (MIW) è stato creato in occasione del lavoro di Dantcheva [8] del 2013. Il database è stato realizzato prelevando delle foto direttamente da internet (per questo il nome wild del database) dove i soggetti vengono ritratti con e senza il makeup (figura 2.11). Siccome le foto trovate in internet ritraevano persone in posa (figura 2.12), le immagini sono state pre-processate ritagliando l’immagine attorno al volto e ridimensionandole a 130 x 150.

MIW contiene foto di 125 soggetti con 1 o 2 immagini per soggetto. In totale il database contiene 154 immagini, di cui 77 senza makeup e 77 con makeup. Ciò significa che a differenza dei precedenti database, MIW non contiene la versione senza trucco e con trucco per ogni soggetto. Questo database contiene foto acquisite in un ambiente estremamente non controllato.

In [8] YMU è stato usato per la fase di *training*, mentre MIW è stato usato durante la fase di test per mettere alla prova il metodo di rilevamento del makeup proposto in quello studio su di un volto in condizioni generiche.



Figura 2.12: Immagini del MIW non pre-processate.

2.3.4 Facial Cosmetics Database



Figura 2.13: Immagini di riferimento e makeup

Facial Cosmetics Database [12] (FCD) è una collezione di immagini acquisite da vari tutorial di makeup presenti su Youtube o su siti web di cosmetica. Questo database è stato creato da Eckert per testare i suoi studi [11]. Le immagini sono state raccolte da vari video tutorial per cercare di riprodurre condizioni di luminosità e posizione del mondo reale (come ad esempio telecamere di sorveglianza).

I volti presenti in FCD appartengono a 50 soggetti, per lo più a donne di razza caucasica tra i 20 e i 30 anni, ma sono presenti, anche se in misura inferiore, rappresentanti del mondo ispanico, asiatico e afroamericano. Per ogni persona viene presa una foto di riferimento presa in diverse condizioni di luminosità,

pelle e concitura e una o più immagini che rappresentano una graduale trasformazione (partendo da senza makeup” a “makeup pesante”).

In questo database sono presenti le immagini sia nella loro versione originale (Figura 2.13) sia in versione ritagliata ed allineata (figura 2.14). Le dimensioni delle immagini ritagliate sono 150 x 140 pixel per la faccia, 60 x 140 per gli occhi e 40 x 70 per la bocca. Sono presenti 389 immagini di cui 50 di riferimento e 339 appartenenti alla graduale applicazione di makeup.

Le 389 sono suddivise come segue:

- 109: no makeup;
- 147: makeup leggero;
- 54: makeup intermedio;
- 79: makeup pesante



Figura 2.14: campioni di immagini di riferimento e makeup ritagliate di FCD

2.3.5 FAcE Makeup database



Figura 2.15: Campione di immagini da FAM.

FAcE Makeup database è stato creato in [16]. Le immagini contenute nel database sono immagini, di personaggi pubblici e celebrità, acquisite da internet (figura 2.15).

In totale sono presenti immagini di 519 soggetti, 222 uomini e 297 donne. Per ogni soggetto è presente un'immagine con makeup e una senza.

2.4 Impatto del makeup sul riconoscimento del volto

Ueda e Koyama [35] hanno condotto degli studi inerenti all'abilità umana di riconoscimento del volti anche in presenza di makeup. Da questi studi hanno concluso che l'abilità di una persona di riconoscere un individuo incrementa se quest'ultimo ha applicato sul suo volto del trucco leggero. Questo perché l'uso di trucco in maniera leggera accentua le caratteristiche distintive proprie

dell'individuo ed è proprio questo fattore che permette di contraddistinguere una persona dagli altri.

Per quanto riguarda il makeup pesante, invece, la situazione cambia completamente. Applicando su di un volto questo tipo di makeup, si cambia drasticamente la percezione delle dimensioni e della simmetria di occhi e bocca andando così a nascondere le caratteristiche che contraddistinguono un volto umano.

Di seguito verranno analizzati i risultati degli studi di Dantcheva [10] e Eckert [11] per verificare se un sistema di riconoscimento facciale che non utilizza tecniche in grado di contrastare l'effetto del makeup possa, come succede per gli esseri umani, essere ingannato. Gli esperimenti presentati usano tecniche, come LBP, che non richiedono fasi di apprendimento per l'estrazione delle feature e per l'analisi. Questo gli attribuisce una considerevole efficienza computazionale. Per il primo e secondo esperimento useremo tre tecniche diffuse per il riconoscimento del volto (i filtri di Gabor [19] (figura 2.16-c), il Local Binary Pattern [34] (figura 2.16-d) e il software commerciale Verilook Face Toolkit [37], mentre per il terzo ci limiteremo all'uso di LBP.

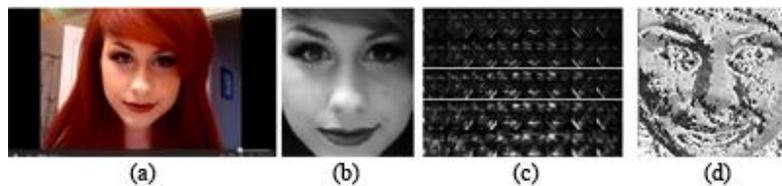


Figura 2.16: (a) Makeup preso da YMU, (b) immagine pre-processata, (c) immagine dopo Gabor. (d) immagine codificata con LBP

2.4.1 Esperimento 1 su YMU

In questo esperimento YMU database viene usato per analizzare di quanto possono diminuire le prestazioni del riconoscimento del volto applicate ad immagini di makeup reale.

Nel database (vedi paragrafo 2.3.1) sono presenti per ogni soggetto 2 immagini con makeup e due senza, che possono essere identificate con M_1 e M_2 per le due con trucco e N_1 e N_2 per le due foto senza. A questo punto sono state sottoposte le immagini a confronto secondo il seguente schema:

- N_1 con N_2 (immagini prima dell'applicazione del makeup)
- M_1 con M_2 (immagini dopo l'applicazione del makeup)
- N_1 con M_2 (un'immagine non ha il makeup mentre l'altra si)

Nella tabella 3 sono riassunti i risultati degli esperimenti. I valori percentuali riportati si riferiscono all'Equal Error Rates (EER_s), ovvero la percentuale di errore che i test hanno generato.

	M vs M	N vs M	N vs N
Gabor	11,59 %	21,47 %	7,01 %
LBP	9,41 %	18,71 %	6,50 %
Verilook	13,55 %	23,68 %	10,85 %

Tabella 3: Risultati dei confronti tra le varie immagini di YMU senza (N) e con (M) makeup

I risultati indicano che la percentuale maggiore di errore si verifica quando viene confrontata un'immagine N con una M dove l'EER passa da un 18,71% (LBP) ad un 23,68% (Verilook). L'EER durante il confronto N con N, invece oscilla da 6,50% (LBP) a 10,85% (Verilook).

Dai dati ottenuti è facile capire come l'applicazione di cosmetici su di un volto può mettere a dura prova un sistema di riconoscimento facciale che non adotti metodi appositi. Il software commerciale Verilook ottiene risultati peggiori tra i tre metodi di confronto in quanto non offre opzioni di ottimizzazione.

	(a) vs (b) (\mathcal{N}_1 vs \mathcal{N}_2)	(a) vs (c) (\mathcal{N}_1 vs \mathcal{M}_1)	(a) vs (d) (\mathcal{N}_1 vs \mathcal{M}_2)
Gabor	0.404	0.573	0.527
LBP	0.447	0.505	0.487
Verilook	0	0.399	0.45

Figura 2.17: Risultati di un campione del database YMU. (a) e (b) soggetto senza makeup, (c) e(d) soggetto con makeup.

In figura 2.17 sono riportati i risultati di un singolo campione preso dal database YMU riportando i punteggi di similarità dei confronti. Come ci si poteva aspettare dai risultati della tabella 3, anche in questo caso si ha un peggioramento di efficienza nel sistema di riconoscimento dopo l'applicazione del makeup indipendentemente dal sistema utilizzato.

La figura 2.18 mostra la distribuzione dei punteggi utilizzando LBP analizzando tre casi: M_1 con M_2 , M con N e N_1 con N_2 . Si può notare come l'applicazione del makeup allontana i valori dalla similarità. Il punteggio peggiore di confronto si ha quando si mettono a confronto le immagini con makeup (M) con quelle senza (N). Alla luce di questi risultati possiamo concludere che l'applicazione di makeup ha il potenziale di ridurre la precisione di confronto degli algoritmi classici di riconoscimento del volto. Si può intuire quanto l'uso di makeup possa essere usato a scopi criminosi aumentando i rischi nel campo della sicurezza rendendo necessario lo studio di nuove tecniche che inibiscano gli effetti potenzialmente negativi dell'uso del trucco.

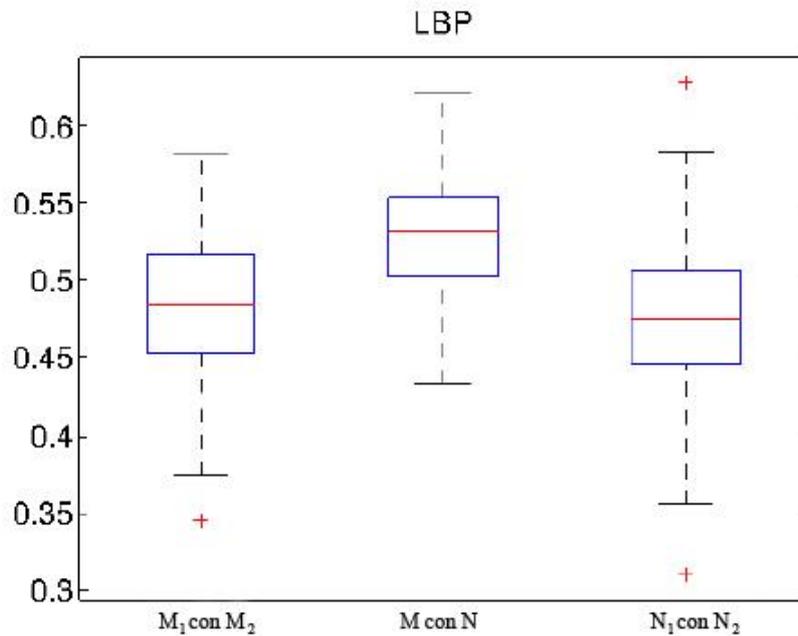


Figura 2.18: Distribuzione dei punteggi di similarità nei casi M_1 con M_2 , M con N e N_1 con N_2 .

2.4.2 Esperimento 2 su VMU

In questo esperimento il database VMU viene usato per determinare quale zona del volto (labbra, occhi o intero volto) abbia il maggior impatto sul riconoscimento del volto dopo l'applicazione di cosmetici. Come descritto nel paragrafo 2.3.2, il Virtual Makeup database contiene immagini di 51 soggetti e per ogni soggetto sono presenti quattro foto (una senza makeup e tre con il makeup applicato solo nelle labbra, solo negli occhi o nell'intero volto). Come per l'esperimento 1 useremo la lettera N per riferirci all'immagine priva di makeup, mentre useremo L per indicare quella nella quale il trucco è presente solo sulle labbra, E per gli occhi e F per l'intero volto.

Le immagini appena descritte sono state poi sottoposte a confronto secondo il seguente schema:

- N con L.
- N con E.
- N con F.

Come per il paragrafo 2.4.2 riportiamo di seguito la tabella con i valori in percentuale degli EER.

	N vs E	N vs L	N vs F
Gabor	8,56 %	8,15 %	11,38 %
LBP	4,32 %	3,43 %	4,79 %
Verilook	25,00 %	4,79 %	30,00 %

Tabella 4: Risultati dei confronti tra le varie immagini di VMU senza makeup (N) con le immagini con trucco (solo occhi E, solo labbra L e trucco completo F)

Grazie ai risultati riportati in tabella 4 è facile dedurre che l'applicazione di makeup sulle labbra influisce in maniera minore sul risultato del confronto (oscillando da un 3,43% con LBP a un 8,15% con Gabor), rispetto all'applicazione di trucco nell'area degli occhi (oscillando da un 4,32% con LBP a un 25% con Verilook).

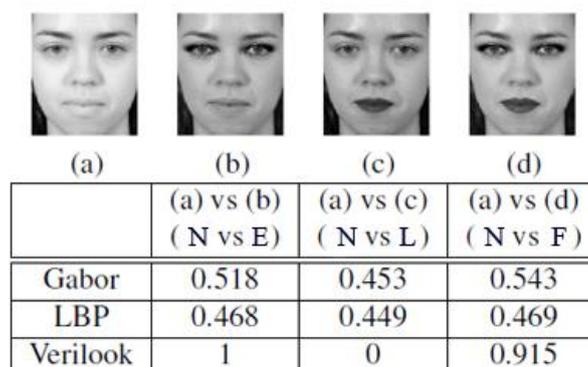


Figura 2.19: Risultati di un campione del database VMU. (a) soggetto senza makeup, (b-d) soggetto con makeup (occhi E, labbra L e completo F).

In figura 2.19 viene mostrato un esempio campione del database VMU con le varie versioni del volto del soggetto scelto: senza makeup (a), con i soli occhi truccati (b), le sole labbra truccate (c) e con il makeup completo (d). E' stato poi messo a confronto il volto senza trucco con le altre tre immagini usando le tre tecniche di riconoscimento ed i dati ottenuti sono stati riportati nella tabella.

La figura 2.19, assieme alla figura 2.20, mostra come il match dei volti senza makeup e con labbra truccate generi il più basso punteggio, mentre il confronto tra volto senza makeup e makeup completo genera il valore più alto.

Questi risultati ci danno un'idea di quanto incida la zona degli occhi sul risultato di un corretto riconoscimento del volto.

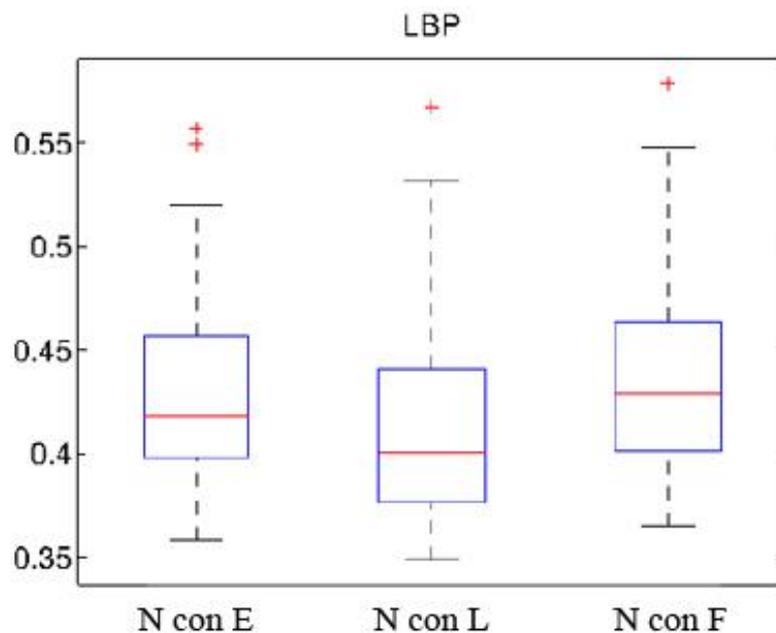


Figura 2.20: Distribuzione dei punteggi di similarità nei casi N con E, N con L e N con F

2.4.3 Esperimento 3 su FCD

Per il terzo esperimento è stato usato l'algoritmo LBP applicato al database FCD (paragrafo 2.3.4). Questo esperimento è diviso in due parti: nella prima parte verranno usate come immagini di riferimento per i confronti, le immagini senza makeup non facenti parte della serie della sequenza di trasformazione del makeup. Nella seconda parte, invece verranno prese come immagini di riferimento quelle con il trucco di pesantezza intermedia. Per verificare quanto preciso sia il risultato dei test verrà usata la percentuale di identificazione (IDR).

Immagini senza makeup come riferimento.

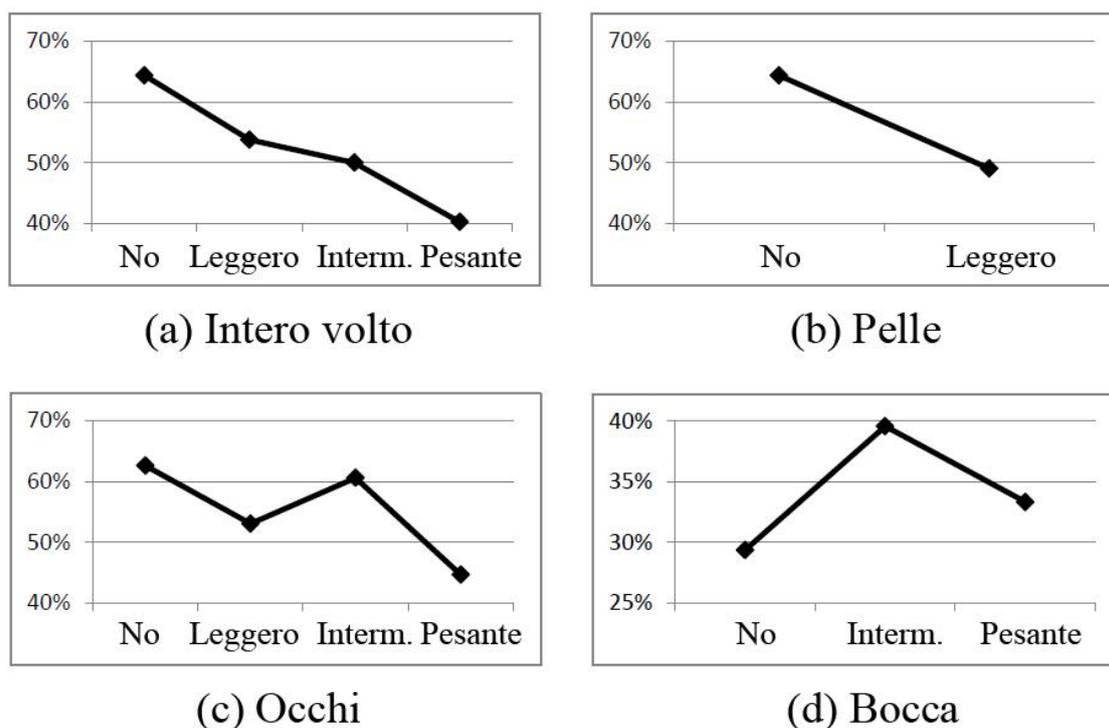


Figura 2.21: IDR prima parte esperimento 3.

Come era facile prevedere, la capacità di un sistema di riconoscimento diminuisce al crescere della quantità di makeup applicato (figura 2.21). Guardando il grafico dell'intero volto (figura 2.21-a) si nota che l'IDR decresce in maniera lineare passando dalla categoria No, Leggero e Pesante, mentre

contrariamente a quello che ci si aspetterebbe, in corrispondenza della categoria Intermedio non si ha una diminuzione lineare, ma decresce meno rispetto all'andamento generale.

L'andamento di questo grafico sembra confermare i due esperimenti precedenti. In corrispondenza della categoria intermedio si ha un calo di IDR inferiore perché è un tipo di trucco molto visibile, ma al contempo che non va ad alterare quelle che sono le caratteristiche proprie di una persona. Al contrario il makeup di categoria pesante va ad alienare un volto eliminando la percezione delle caratteristiche fondamentali di ogni individuo, mascherando i difetti ed è per questo motivo che in questo caso abbiamo un IDR di solo 40%.

Nel grafico (b) abbiamo solamente la categoria No e Leggero perché il database non è abbastanza grande da contenere tante tipologie anche per il trucco della pelle. Nonostante questo, si nota che anche solo l'applicazione di makeup Leggero alla pelle decrementa di molto l'IDR, questo risultato è dovuto al fatto che il makeup per la pelle uniforma il colore della stessa, e cambia la percezione della forma del viso.

In figura 2.21-c sono rappresentati i dati dei confronti basati sull'area degli occhi e si nota subito che l'andamento generale della curva è lo stesso del grafico (a). L'unico dato degno di essere specificato è che la zona occhi, in corrispondenza della categoria Intermedia, ha circa lo stesso IDR della zona occhi senza makeup in quanto il trucco in questa zona accentua molto di più, rispetto ad altre zone, le caratteristiche dell'occhio, rendendo ben distinta la sua forma. Ancora una volta il makeup applicato in maniera pesante incide negativamente sul riconoscimento mascherando l'occhio.

Infine dal grafico della bocca (figura 2.21-d) si vede come la curva è molto più bassa rispetto a quella delle altre zone del volto. L'IDR della zona della bocca è il più basso perché il riconoscimento applicato alle labbra è molto sensibile al cambio di espressione. Questo porta alla conclusione che la bocca è la parte che meno contribuisce durante il riconoscimento di un volto. Comunque contrariamente alle altre zone del volto, l'applicazione del makeup all'area della bocca, contribuisce in maniera positiva al riconoscimento. E' facile intuire che questo risultato è dovuto al fatto che il trucco sottolinea i contorni delle labbra e

ne aumenta il contrasto. Come per le altre zone il trucco intermedio conduce ad un IDR più alto. In questo caso il trucco pesante diminuisce l'IDR rispetto al trucco intermedio, ma comunque rimane più alto rispetto all'assenza totale di makeup. Come si vedeva da figura 2.3, non è presente la categoria leggero per la zona della bocca.

In definitiva la zona degli occhi sembra essere la parte cruciale durante la fase di riconoscimento di un volto [23], il makeup intermedio favorisce il riconoscimento di un volto, l'applicazione di makeup alla pelle riduce la riconoscibilità di un individuo e il makeup applicato alla bocca produce un effetto positivo sul riconoscimento anche se in misura poco significativa. La domanda a cui tentiamo di rispondere con la seconda parte di esperimento è la seguente: può un'immagine con makeup intermedio, presa come riferimento di partenza, aumentare il riconoscimento di un volto?

Immagini con makeup di categoria Intermedio come riferimento.

Il test è stato condotto usando 23 immagini con makeup intermedio come immagini di riferimento per i confronti. A differenza della prima parte del test, dove l'immagine usata era prelevata in altre condizioni rispetto alla serie di makeup, in questo caso l'immagine di riferimento è acquisita nello stesso ambiente delle altre. Questo test risulta quindi meno realistico rispetto al precedente, ed essendo l'immagine di riferimento facente parte della categoria Intermedia è logico che in corrispondenza di questa categoria si abbia un match del 100% (figura 2.22).

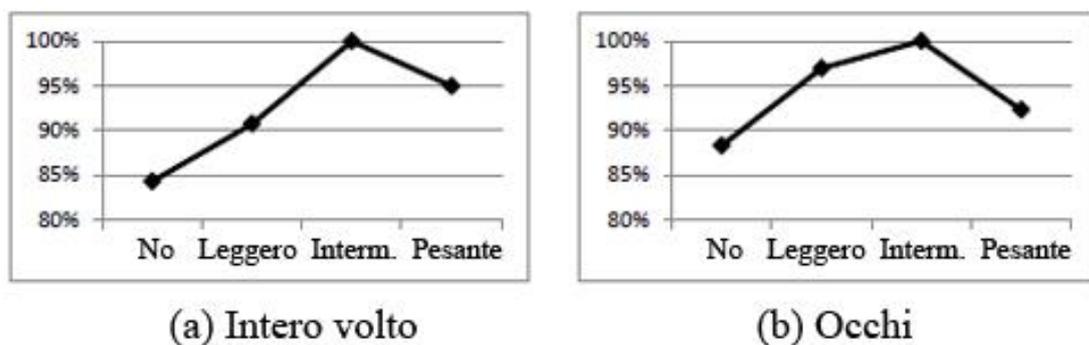


Figura 2.22: IDR seconda parte esperimento 3

E' evidente che le curve (figura 2.22-a e 2.22-b) sono più alte rispetto al precedente test (figura 2.21): questi risultati sono dovuti dal fatto che come già detto le immagini sono prese dalla stessa serie di quelle di riferimento, ma anche perché grazie all'effetto intermedio del makeup, le caratteristiche proprie di un volto sono accentuate e quindi più riconoscibili da un sistema di riconoscimento facciale. Aumentando il makeup si ha quindi un aumento quasi lineare della curva IDR, per poi decrescere in corrispondenza della categoria pesante. Anche in questa seconda parte del test la zona degli occhi si dimostra essere la parte che maggiormente incide durante il riconoscimento di un volto.

2.4.4 Conclusioni

Alla luce dei dati ottenuti con questi tre esperimenti possiamo comprendere l'effetto che ha l'uso del makeup su di un sistema di riconoscimento del volto. Siccome il trucco altera i colori di un volto, usando un classico sistema di confronto basato su delle immagini in scala di grigio, le performance con la presenza di makeup diminuiscono. Sia usando il database realistico YMU che quello virtuale VMU otteniamo dei risultati che indicano un calo di prestazioni quando un volto senza makeup viene confrontato con uno con makeup. Questo è un altro indizio dell'effetto negativo del trucco.

In tutti e tre gli esperimenti risulta chiaro che la zona che maggiormente è soggetta ad alterazione con l'applicazione di makeup è l'area degli occhi.

Dall'esperimento 3 si apprende sorprendentemente che usare come immagini di riferimento delle foto di individui con makeup può aiutare un sistema ad ottenere un migliore punteggio.

Siccome l'uso di makeup è diventato di uso comune in quasi tutto il mondo, è necessario pensare ad algoritmi che contrastino gli effetti negativi provocati dall'uso del trucco. Questo per contrastare l'uso criminale che si potrebbe fare del makeup, che andrebbe a compromettere la nostra sicurezza.

Nel prossimo capitolo verranno elencati vari algoritmi per contrastare gli effetti negativi del makeup sul riconoscimento facciale.

Capitolo 3

Riconoscimento del volto in presenza di makeup

Dopo aver analizzato i dati raccolti dagli esperimenti riportati nel capitolo 2 risulta chiaro, per garantire una pubblica sicurezza, quanto sia necessario pensare ad un sistema di riconoscimento del volto che, diversamente da quello standard (capitolo 1), non sia sensibile agli effetti negativi del makeup.

In questo capitolo verranno raccolti gli algoritmi studiati negli ultimi anni che aiutano a migliorare il riconoscimento di un volto. Il capitolo è strutturato organizzando gli algoritmi in tre categorie:

- Algoritmi in grado di rimuovere il trucco
- Algoritmi in grado di rilevare il trucco
- Algoritmi di confronto robusti al problema del makeup

Dopo aver presentato gli algoritmi verranno condotti degli esperimenti di riconoscimento del volto usando dei software commerciali per testarne la loro efficacia. Infine tutti i risultati ottenuti verranno messi a confronto.

3.1 Algoritmi in grado di rilevare il trucco

Un'altra categoria di algoritmi molto importanti è quella degli algoritmi in grado di rilevare la presenza o assenza di trucco in un volto. Sono algoritmi importanti in un processo di riconoscimento del volto in quanto in caso di presenza di trucco si possono adottare degli accorgimenti per pre-processare le immagini prima di confrontarle. Diversamente se non è presente del makeup è sufficiente utilizzare un semplice sistema di rilevamento del trucco che non preveda ulteriori accorgimenti.

3.1.1 Algoritmo basato su colore, texture e forma

Il seguente metodo estrapola dalle immagini un vettore di *feature* che cattura informazioni relative a forma, *texture* e colore del volto (nella sua interezza e considerando delle sotto aree). Verranno testati poi i risultati con due classificatori: SVM ed Adaboost. Gli esperimenti vengono condotti usando come database YMU e MIW.

Tra i vari spazi dei colori a disposizione (ad esempio RGB, Lab etc.) questo algoritmo utilizza HSV (Hue/Saturation/Value). Questa scelta è dettata dal fatto che le informazioni riguardanti il makeup sono più facilmente riconoscibili con HSV.

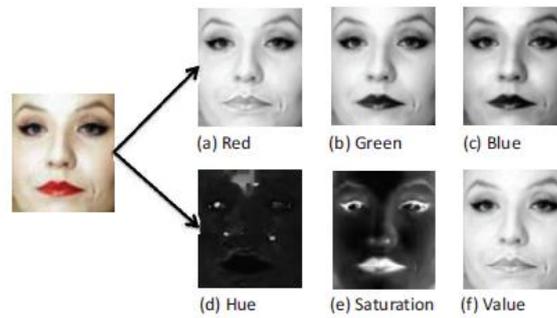


Figura 3.6: Immagine di un volto nel modello di colori RGB e HSV

Come si vede in figura 3.6 il makeup può essere più facilmente rilevato attraverso il canale di saturazione dello spazio (figura 3.6-e). HSV può essere ottenuto tramite una trasformazione non lineare del modello RGB:

$$H = \arctan \left\{ \frac{\sqrt{3}(G - B)}{(R - G) + (R - B)} \right\}; S = 1 - \frac{\min\{R, G, B\}}{V}; V = \frac{(R + G + B)}{3}$$

L'algoritmo prevede vari passaggi: localizzazione del volto, localizzazione dei *landmark*, normalizzazione del volto, estrazione delle aree di interesse (ROI), estrazione delle *feature* e classificazione delle *feature* (figura 3.7).

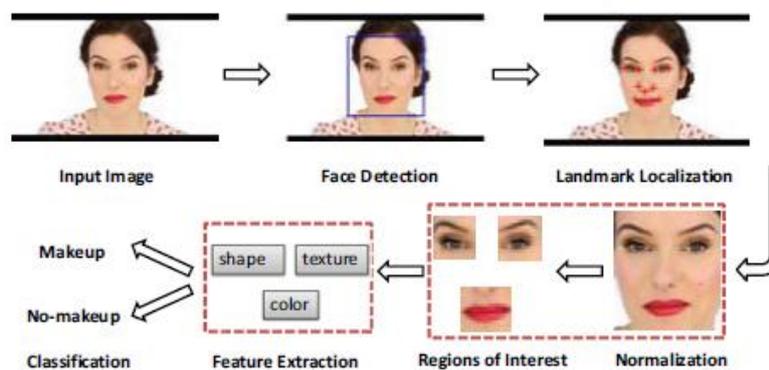


Figura 3.7: Step di classificazione delle feature

Localizzazione e ROI

Per prima cosa viene usato Adaboost per localizzare il volto. I *landmark* delle *feature* vengono inseriti tramite modello GMM (Gaussian Mixture Model). Subito dopo, sapendo dove sono situati gli occhi, è possibile ritagliare e allineare il volto. Il volto viene normalizzato per rimuovere variazioni di posizione e scala, tagliando e ridimensionando l'immagine a 150x130 pixel. Anche le ROI utili all'algoritmo (occhio sinistro, occhio destro e bocca) subiscono gli stessi processi del volto: occhio sinistro e destro 52x52 mentre la bocca 56x62.

Estrazione delle feature

Siccome il makeup va ad alterare i colori in un volto, sono state scelte delle *feature* basate sul colore. Dei filtri di Gabor sono stati usati per estrarre informazioni di forma e *texture*, mentre LBP è impiegato per rilevare micro dettagli nel volto.

A) Descrittore di colore

Ogni ROI viene divisa in 5x5 blocchi non sovrapposti poi per ogni canale vengono calcolati i tre momenti colore (centrale, deviazione standard e indice di simmetria: rispettivamente equazione (a), (b) e (c)).

$$a) \rho = \sum_{x,y} \frac{1}{N} I_{x,y};$$

$$b) \sigma = \sqrt{\frac{1}{N} \sum_{x,y} (I_{x,y} - \rho)^2};$$

$$c) \gamma = \sqrt[3]{\frac{1}{N} \sum_{x,y} (I_{x,y} - \rho)^3}$$

Dove N è il numero totale dei pixel e $I_{x,y}$ indica il pixel alla posizione (x,y) . Così facendo si ottiene un vettore di dimensione 225. Per la faccia invece si divide in 9 blocchi ottenendo un vettore di dimensionalità 81.

B) Descrittore di forma

Vengono usati tre tipi di descrittori di forma per estrapolare delle *feature* aggiuntive. Il primo è basato sui filtri di Gabor ottenendo 40 immagini. Come

primo descrittore vengono calcolati i tre momenti per ognuna delle 40 immagini ottenendo 120 *feature*. Come secondo descrittore viene usato GIST [27]. Il terzo descrittore è basato sulle informazioni dei contorni che vengono sottolineate applicando il makeup.

C) *Descrittore di texture*

Per raccogliere le micro-informazioni viene utilizzato LBP nell'immagine del volto.

La dimensione totale del vettore delle *feature* (colore forma e *texture* ottenuti da ROI e faccia) è 1484 (809 per il volto e 675 per le ROI figura 3.8)

Attribute	Feature	Face-Dim	ROI-Dim
Color	Moments	81	225 × 3
Shape	Gabor	120	-
	GIST	512	
	EOH	37	
Texture	LBP	59	-
Total		809	675

Figura 3.8: dimensione delle feature usate

Classificazione

Per classificare il vettore ottenuto in “makeup” o “senza makeup” vengono usati due classificatori: SVM e Adaboost. SVM cerca un confine lineare che massimizza il margine tra due classi di modelli, mentre Adaboost combina più classificatori più deboli per crearne uno unico più potente.

Esperimenti

Per testare l'algoritmo è stata usata una 5-fold cross validation usando YMU come database (4 per il l'addestramento e una per i test). Il tutto ripetuto 5 volte [8]. Per valutare i risultati viene utilizzato il Classification Rate (CR), ovvero la percentuale di immagini con e senza makeup correttamente classificati.

ROI	SVM (%)	Adaboost (%)
Intero volto	87,25	88,98
Occhio sinistro	81,71	75,72
Occhio destro	80,68	79,89
Bocca	58,94	57,46
Occhio sinistro + Occhio destro + Bocca	87,62	85,83
Volto + Occhio sinistro + Occhio destro + Bocca	91,20	89,94

Tabella 5: Risultati esperimento su YMU

Dalla tabella 5 si nota che l'efficacia dell'algoritmo sulla zona della bocca diminuisce drasticamente. Questo dato potrebbe essere dovuto alla variazione di posizione a cui è soggetta la bocca e all'incapacità dei momenti del colore di catturare l'area che il rossetto ha fatto diventare omogenea. Le *feature* analizzate individualmente restituiscono dei risultati scarsi (momenti colore: 77.62%; Gabor: 62.08%; GIST: 86.79%; EOH: 56.68%; LBP: 50.78%). Fondendo i tre vettori delle ROI il CR sale vertiginosamente a 87,62%, per aumentare ancora fino a 91,20% fondendo tutte le *feature* in un vettore a 1484 dimensioni.

Di seguito sono riportati i risultati di tutte e cinque i test svolti (tabella 6):

Prova	Addestramento	Test	SVM (%)	Adaboost (%)
1	487 (243/244)	113 (57/56)	92,04	91,15
2	473 (237/236)	127 (63/64)	90,55	87,15
3	487 (244/243)	113 (56/57)	91,15	90,27
4	457 (228/229)	143 (72/71)	90,91	89,51
5	796 (248/248)	104 (52/52)	91,35	91,35
Media			91,20	89,94

Tabella 6: Risultati parziali dei 5 esperimenti. I numeri tra parentesi indicano il numero di immagini con makeup e senza makeup

Applicando l'algoritmo al database MIW si sono ottenuti i seguenti CR:

- SVM 95,45 %
- Adaboost 92,21 %

Con questo algoritmo il localizzatore di volti ha fallito in 22 immagini su 154.

3.1.2 Rilevamento del makeup tramite l'uso del Local Fisher Discriminant Analysis

Invece di concentrarsi sull'intero volto, questo algoritmo estrae delle patch locali. Infine vengono confrontati due sistemi di mappatura discriminanti (Marginal Fisher Analysis e Local Fisher Discriminant Analysis).

Patch locali

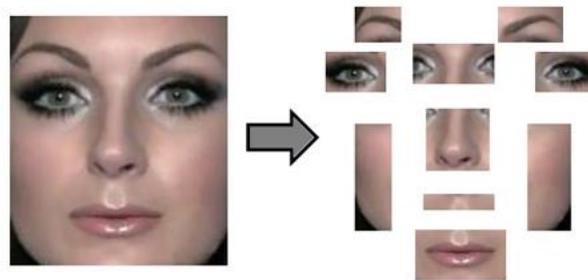


Figura 3.9: Estrazione delle patch locali

Come nel lavoro di Gue, Yan e Wen [15], vengono estratte dall'immagine del volto delle patch locali, 10 per la precisione (figura 3.9). Le immagini vengono allineate, poi le 10 patch locali vengono automaticamente estratte.

Estrazione feature

Per evitare che i dati ridondanti disturbino il classificatore, invece di lavorare direttamente sui pixel dell'immagine, vengono estratte quattro principali *feature*: tonalità del colore, uniformità, *texture* e luminosità.

A) *Tonalità del colore*

In questo algoritmo viene usato il modello di colori RGB in quanto le maggiori modifiche del makeup sono incentrate sulle caratteristiche del colore. Calcolando momento centrale, deviazione standard ed entropia per ognuno dei canali del colore si ottengono, per tutte e 10 le patch estratte, si ottengono 90 feature di tonalità del colore.

B) *Uniformità*

Dopo l'applicazione di makeup la pelle tende ad assumere un colore uniforme nascondendo i difetti. Per misurare questa *feature* vengono di nuovo calcolati momento centrale, deviazione standard e entropia per tutte e 10 le patch ottenendo 30 *feature* di uniformità da ogni immagine in scala di grigi.

C) *Texture*

Il makeup tende a donare alla pelle un aspetto più morbido. Raccogliendo le informazioni sulle *texture* prima e dopo il makeup si contribuisce ad accelerare il processo di rilevamento del trucco. Solitamente per catturare le *texture* di un'immagine viene impiegato l'LBP, ma per questo algoritmo si è preferito utilizzare il WLBP (Weighted Local Binary Pattern) [5].

D) *Luminosità*

Per far sembrare alcune parti del volto più sporgenti, alcuni prodotti di makeup donano più lucentezza alla pelle. Questa luminosità può essere formulata dalla combinazione lineare dei componenti della luce riflessa in ognuno dei canali dello spazio RGB.

Selezione delle patch

Invece di utilizzare tutte le patch, questo algoritmo prevede di selezionare le patch che hanno ottenuto il valore massimo nella rilevazione di makeup per ogni feature.

In tabella 7 vediamo come migliora la precisione passando dalla considerazione dell'intero volto ad un'attenta selezione delle patch.

Feature	Precisione		
	Intero volto	Patch locali	Patch selezionati
Colore	55,64%	77,24%	79,53%
Uniformità	52,50%	73,81%	80,23%
Texture	63,28%	70,68%	72,44%
Luminosità	52,50%	64,50%	76,44%

Tabella 7: Dati di precisione nella rilevazione del makeup usando l'intero volto, le patch locali o una selezione delle patch locali

Dimensionality Reduction

Vengono usati nell'algoritmo due metodi per integrare *feature* sulle patch selezionate vengono valutati due metodi: Marginal Fisher Analysis (MFA) e Local Fisher Discriminant Analysis (L DFA).

Nonostante MFA dia degli ottimi risultati, L DFA identifica la combinazione lineare delle variabili che permette di separare al meglio le diverse classi. L DFA proietta un set di funzioni su di uno spazio di funzioni ridotto mantenendo allo stesso tempo le informazioni relative alla discriminazione tra classi.

Classificazione

Come classificatore viene usato SVM con Radial Basis Function (RBF) come kernel. Anche in questo algoritmo si è adottato lo schema di dividere il database in cinque cartelle, usandone quattro per l'addestramento e una per il test. Il tutto ripetuto cinque volte.

Esperimenti

Per testare l'algoritmo sono stati usati YMU e VMU come database [30].

Ogni *feature* su ogni patch risponde in maniera differente per la rilevazione del trucco. Vediamo ora in dettaglio quale patch aggiungere per ottenere i migliori risultati. Ad ogni prova vengono aggiunte una ad una le patch e viene tenuta quella che garantisce un miglior risultato.

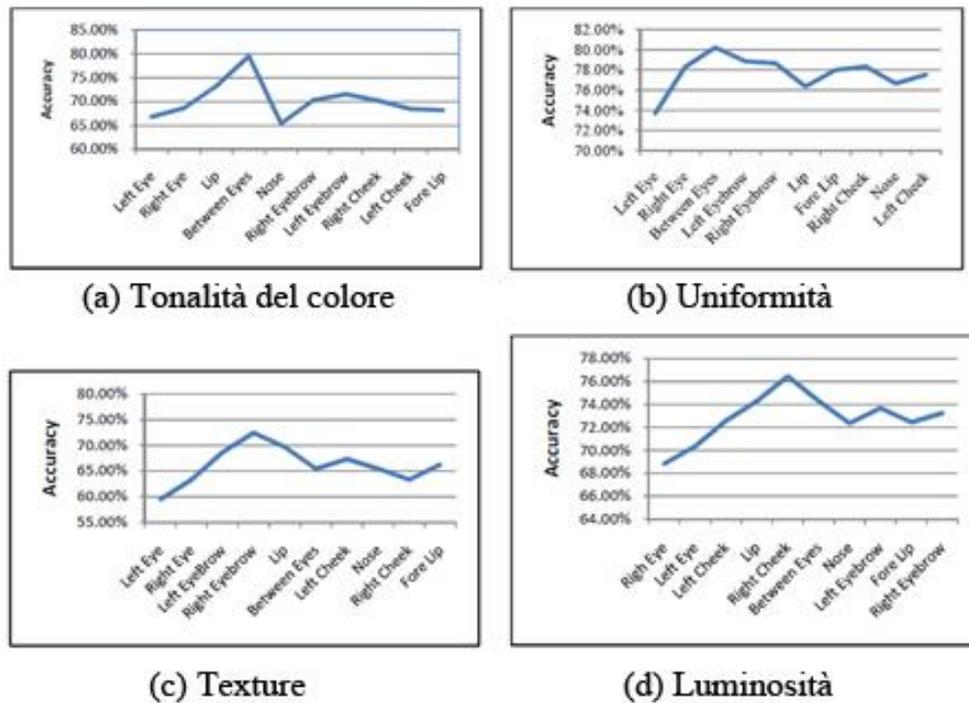


Figura 3.10: Risultati di precisione delle feature per le diverse patch

Per quanto riguarda la tonalità del colore (figura 3.10-a) si raggiunge il valore massimo di precisione nella rilevazione del makeup con l'aggiunta dell'occhio sinistro, l'occhio destro, le labbra e lo spazio tra gli occhi (80%). Aggiungendo un'altra patch il livello di precisione diminuisce. Le primi quattro patch sono state selezionate e mappate con LDFA.

Per l'uniformità della pelle (figura 3.10-b) le patch migliori sono gli occhi e lo spazio tra di essi (80%).

Per le *texture* (figura 3.10-c) il valore massimo viene raggiunto con l'uso degli occhi e delle sopracciglia (72%).

Per la luminosità (figura 3.10-d) i valori più alti sono ottenuti in corrispondenza dell'occhio destro e della guancia destra. Questo risultato è in contrasto con le aspettative, secondo cui le patch maggiormente interessate per questa *feature* dovrebbero essere guance e labbra. Questo risultato inaspettato potrebbe essere dovuto dalla direzione della luce che colpisce il volto. Il miglior risultato (76%) si raggiunge con l'aggiunta di occhi, guance e labbra.

Di seguito mostriamo i dati effettivi per ogni patch relativi ad ogni *feature*.

Patch locali	Precisione			
	Tonalità colore	Uniformità	Texture	Luminosità
Occhio sinistro	66,83%	73,66%	59,44%	60,44%
Occhio destro	65,83%	72,66%	53,25%	68,83%
Labbra	63,16%	60,00%	55,50%	55,50%
Sopracciglio sinistro	60,46%	55,00%	58,26%	52,44%
Sopracciglio destro	62,58%	53,33%	55,50%	62,78%
Tra gli occhi	61,33%	67,00%	57,66%	59,64%
Naso	59,75%	48,50%	51,76%	53,87%
Sopra al labbro	48,33%	46,66%	53,12%	50,50%
Guancia sinistra	51,16%	53,16%	50,00%	57,81%
Guancia destra	55,00%	50,16%	50,00%	63,29%

Tabella 8: Dati delle diverse feature per ogni patch

Oltre ad aver selezionato le patch locali che garantiscono i risultati migliori, viene creata una mappatura delle differenti combinazioni di *feature* e patch per ottenere informazioni maggiormente discriminanti. Di seguito riportiamo i risultati ottenuti applicando prima la mappatura MFA e poi LFDA.

Combinazioni di feature dopo la selezione locale delle patch	MFA	LDFA
Uniformità + Luminosità	70,27%	81,66%
Colore	71,24%	83,33%
Colore + Luminosità	72,85%	84,50%
Colore + Texture	76,36%	85,98%
Colore + Uniformità	71,50%	87,50%
Colore + Uniformità + Luminosità	73,25%	89,16%
Colore + Uniformità + Texture + Luminosità	81,24%	95,03%

Tabella 9: Diverse combinazioni di feature con MFA e LDFA a confronto

La miglior precisione ottenuta è data dall'uso di LDFA utilizzando tutte e quattro le *feature* ed è pari a 95,03%.

Utilizzando le tecniche descritte in questo algoritmo, con i dati ottenuti, risulta chiaro che si ottiene una maggior precisione usando patch locali selezionate piuttosto che l'intero volto. In più per trovare maggiori correlazioni tra due immagini con e senza makeup e ridurre le dimensioni delle *feature* si è visto che LDFA ottiene un ottimo risultato applicato su YMU.

3.1.3 Algoritmo basato sull'analisi di forma e texture (LGBP + HOG)

Con il seguente metodo (messo a punto da Kose, Apriville e Dugelay [18]), data l'immagine di un volto, vengono prima estratte le caratteristiche (*feature*) basate su forma e *texture* della faccia. Queste caratteristiche vengono estratte usando il Local Gabor Binary Pattern (LGBP) [44] come descrittore di trama e Histogram of Gradients (HOG) come descrittore di forma. I risultati vengono poi utilizzati da un classificatore per individuare l'eventuale presenza di makeup (figura 3.11). Per testare questo algoritmo sono stati usati tre database: YMU, MIW e FCD, mentre come classificatori sono stati impiegati la Support Vector Machine (SVM) [7] e Alligator.

Diversamente da altri algoritmi, questo non utilizza le informazioni sul colore per determinare la presenza di makeup.

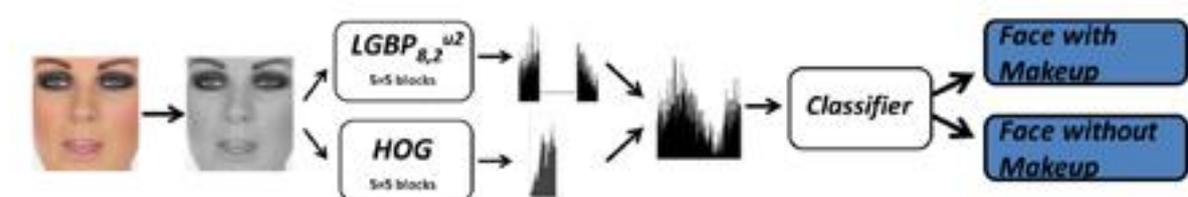


Figura 3.11: Processo di rilevamento del makeup

Estrazione delle feature

In molti ambiti (rilevamento di persone, verifica di firme, rilevamento di oggetti, etc.) vengono usati HOG e LBP assieme. In questo algoritmo si testa la sostituzione di LBP con LGBP, ovvero l'uso di LBP su di un'immagine ottenuta tramite Gabor.

A) LGBP

Siccome le forme e le *texture* di un volto subiscono gli effetti del makeup, per prima cosa l'immagine in input viene pre-processata per ottenere più immagini GMP (Gabor Magnitude Pictures) in un dominio di frequenza.

Queste immagini sono ottenute applicando dei filtri di Gabor multi-scala e multi orientazione. Ogni immagine GMP viene poi convertita ad una mappa LGBP che viene poi divisa in regioni rettangolari non sovrapposte di 5x5. Gli istogrammi LGBP ottenuti da tutte le mappe sono poi concatenati per creare un unico istogramma finale che definisce il modello del volto.

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,v}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{\mu,v}z} - e^{-\frac{\sigma^2}{2}} \right]$$

Nell'equazione viene definito il filtro di Gabor, dove μ e ν rappresentano orientamento e scala dei filtri di Gabor, $K_{\mu,\nu}$ rappresenta il vettore onda, definito come $K_{\nu}e^{i\phi_{\mu}}$, dove $K_{\nu}=K_{\max}/\lambda^{\nu}$ e $\phi_{\mu}=\pi\mu/8\lambda$.

La rappresentazione di Gabor del volto è ottenuto tramite la convoluzione dell'immagine con i filtri di Gabor. Sia $f(x,y)$ l'immagine del volto, la sua convoluzione con i filtri di Gabor ($\psi_{\mu,\nu}(z)$) è definita come segue:

$$G_{\psi_f}(x, y, \mu, \nu) = f(x, y) * \psi_{\mu,\nu}(z)$$

Dove $*$ definisce l'operatore della convoluzione. Per queste operazioni $\nu \in \{0, \dots, 4\}$ e $\mu \in \{0, \dots, 7\}$. Alla fine del processo si ottengono 40 GMP. I valori vengono poi codificati con LBP.

Alla fine il numero totale di *feature* estratte da un'immagine è pari a $59 \times 5 \times 5 \times 40 = 59000$.

B) HOG

In questo algoritmo le celle si sovrappongono per metà della loro area, questo significa che ogni cella contribuisce più di una volta alla creazione del vettore finale delle *feature*. Per tenere conto delle variazioni di illuminazione e contrasto, i valori dei gradienti sono normalizzati localmente per ogni cella. Come suggerito in [21], sono stati usati nove celle rettangolari e per ogni cella sono stati calcolati nove rettangoli di istogramma. I nove istogrammi, con i loro nove rettangoli, sono stati poi concatenati per creare un vettore di *feature* di 81 dimensioni. La tecnica HOG è stata applicata su 5×5 aree non sovrapposte. Il numero totale di *feature* estratte da un'immagine risulta quindi essere $81 \times 5 \times 5 = 2025$.

C) Fusione

I valori delle *feature* di HOG sono tra 0 e 1 a causa della normalizzazione presente nella sua implementazione. Per questo motivo è stata applicata la stessa normalizzazione anche per le *feature* di LGBP. Dopo la normalizzazione le *feature* estratte con LGBP e HOG sono state concatenate.

Per valutare l'efficacia dell'algoritmo proposto vengono utilizzati due classificatori: SVM e Alligator. Alligator è un tool che contiene al suo interno SVM, K-NN, epsilon-cluster e altri.

Esperimento 1

Il primo esperimento è stato svolto utilizzando il database YMU [18]. Per poter effettuare dei confronti l'esperimento è stato svolto seguendo lo schema degli esperimenti di Dantcheva [8]: YMU diviso in 5 cartelle (30 soggetti circa in ognuna), di cui 4 per la fase di training e una per il test. Il tutto è stato ripetuto 5 volte.

Per questo esperimento viene utilizzato solamente SVM come classificatore e l'efficienza del test viene misurato in base al Classification Rate (CR), ovvero la percentuale di immagini con e senza makeup correttamente classificati.

Prova	Addestramento	Test	SVM (%)
1	480 (240/240)	120 (60/60)	100
2	480 (240/240)	120 (60/60)	93,33
3	480 (240/240)	120 (60/60)	100
4	480 (240/240)	120 (60/60)	100
5	480 (240/240)	120 (60/60)	99,17
Media			98,50

Tabella 10: Risultati test 1 su YMU. I numeri tra parentesi indicano il numero di immagini con makeup e senza makeup

Nel paragrafo 3.1.1 l'algoritmo analizzato otteneva (utilizzando il medesimo schema) una media del 91,20%, mentre con questo algoritmo si riesce ad ottenere un risultato del 98,5% usando lo stesso classificatore.

Per meglio comprendere l'efficienza della fusione di LGBP con HOG sono stati condotti gli stessi test usando solo le *feature* dei due descrittori singolarmente. I risultati ottenuti sono di una media del 89,5% per HOG e 82,5% per LGBP. L'incremento che si ottiene dalla fusione dei due descrittori è la prova che HOG (descrittore di forma) e LGBP (descrittore di texture) sono complementari.

Esperimento 2

Per il secondo esperimento l'algoritmo è stato addestrato attraverso l'uso di 300 volti con makeup e 300 volti senza makeup presi da YMU, per poi essere testato su MIW (77 volti con makeup e 77 senza makeup) [18].

Anche questo esperimento è stato svolto seguendo lo schema di un esperimento riportato nel paragrafo 3.1.1 così da poter confrontare ancora una volta i risultati. L'algoritmo utilizzato nel paragrafo 3.1.1 aveva ottenuto un CR di 95,45%, usando SVM, e 92,21% usando Adaboost come classificatore. Sempre durante quell'esperimento era stato ottenuto un 1% come FPR e un TPR di 93,51% (SVM) e 84,42% (Adaboost).

Con l’algoritmo descritto in questo paragrafo i dati ottenuti sono di un CR di 99,35% (SVM), con 1% di FPR e 97,70% TPR. I risultati comparati sembrano provare un netto miglioramento di questo algoritmo rispetto al precedente, raggiungendo quasi la perfezione. Delle 154 immagini, questo algoritmo ha fallito nella classificazione con una sola immagine (figura 3.12).



Figura 3.12: L’immagine classificata in maniera scorretta

Guardando la figura 3.15 si può supporre che l’errata classificazione della stessa sia dovuta alla posa non frontale del volto nell’immagine.

Questo esperimento sottolinea la capacità generale del seguente algoritmo essendo esso costruito, nelle fasi di training e del test, su due database diversi.

Esperimento 3

Nel terzo esperimento è stato applicato l’algoritmo sul database FCD (usando lo stesso schema delle 5 cartelle dell’esperimento 1) al fine di testare l’efficienza anche su altre tipologie di database [18].

Prova	Addestramento	Test	SVM (%)	Alligator (%)
1	270 (189/81)	115 (87/28)	90,43	100
2	325 (236/89)	60 (40/20)	83,33	100
3	324 (235/89)	61 (41/20)	96,72	100
4	319 (230/89)	66 (46/20)	87,88	96,05
5	302 (214/88)	83 (62/21)	87,95	97,60
Media			89,26	98,73

Tabella 11: Risultati test 3 su FCD. I numeri tra parentesi indicano il numero di immagini con makeup e senza makeup

Dopo l'estrazione delle *feature*, la fase di addestramento per SVM impiega 16 secondi, mentre per Alligator impiega meno di un minuto. In base ai risultati ottenuti, questo algoritmo, dimostra di essere preciso anche se sottoposto a diversi database. Il valore medio del CR in questo esperimento usando SVM (89,26%) è inferiore rispetto a quello ottenuto nell'esperimento 1 (98,50%), questo dato dimostra che FCD è una sfida maggiore per un sistema di riconoscimento facciale rispetto a YMU.

L'uso di Alligator come classificatore, invece di SVM, mostra quanto sia efficiente la fusione di diversi classificatori (aumentando da 89,26% a 98,73%).

Esperimento 4

Nei precedenti esperimenti, l'algoritmo era messo alla prova come classificatore binario. In quest'ultima prova è stato testato [18] se l'algoritmo è in grado di effettuare una classificazione multipla. Come già detto al paragrafo 2.3.4, FCD contiene volti con diversi gradi di makeup (no, leggero, intermedio e pesante). Siccome le immagini a nostra disposizione che rappresentano la categoria intermedia sono poche (52), esse vengono fuse assieme alla categoria pesante, quindi le classi sottoposte a test saranno no, leggero e pesante. Le immagini dei primi 40 soggetti sono state usate per la fase di training delle tre categorie, mentre le rimanenti sono usate per la fase di test.

Su 83 campioni totali, 60 sono stati classificati correttamente (72,29%).

Le 23 immagini classificate in maniera scorretta si dividono come segue:

- 2 su 21 per la categoria No
- 5 su 27 per la categoria Leggero
- 16 su 35 per la categoria Pesante

Come ci si poteva aspettare il seguente algoritmo perde efficacia rispetto alla versione binaria, ma si può concludere che rimane un buon algoritmo anche in caso di classificazione multipla.

Con questi risultati l'algoritmo sembra perfetto per essere usato in una fase di pre-elaborazione delle immagini in un sistema di riconoscimento del volto.

3.2 Algoritmi in grado di rimuovere il trucco

Questo algoritmo tenta di eliminare, tramite tecniche, gli effetti del makeup che vanno ad inficiare i risultati di un sistema di riconoscimento del volto.

3.2.1 Algoritmo basato sull'algebra di Clifford

L'algoritmo presentato utilizza i concetti di colore e riconoscimento delle immagini multispettrali (ovvero quelle immagini multicolore e multicomponente). Le immagini RGB sono un classico esempio di immagini a colori composte da tre componenti separate: rosso (R), verde (G) e blu (B).

Gli animali con differenti storie evolutive hanno sistemi visivi con diverse dimensionalità. Il cervello umano usa tre canali (RGB) per riconoscere le immagini a colori. I primati invece possiedono un sistema visivo dicromatico e tricromatico e possono usare diversi canali a due o 3 dimensioni per gli stessi scopi. Gli esseri non primati hanno dei sistemi visivi monocromatici e dicromatici. Le immagini di questi sistemi sono funzioni a valori reali o complessi.

Nell'approccio classico, ogni pixel multicolore, è associato ad un punto di uno spazio vettoriale multicolore a k -dimensioni (per le immagini a colori si ha uno spazio vettoriale 3D RGB).

In questo paragrafo mostreremo che con l'algebra di Clifford si possono sfruttare queste nozioni per rimuovere il makeup da un volto in maniera efficiente.

Algebra di Clifford

L'algebra Geometrica è uno strumento matematico formale che offre numerosi strumenti per la risoluzione di problemi specificatamente geometrici unificando ed estendendo algebra lineare, geometria differenziale e calcolo vettoriale.

L'algebra di Clifford è un tipo particolare di algebra Geometrica. Il nucleo fondamentale su cui si basa l'algebra di Clifford è il prodotto di Clifford che è un prodotto vettoriali di due vettori a e b (figura 3.1). Questo prodotto è composto da due parti: una scalare data dal prodotto scalare $a \cdot b$ e una bivettoriale data dal prodotto vettoriale $a \wedge b$.

$$ab = a \cdot b + a \wedge b$$

Come già detto questo prodotto unifica prodotto scalare e vettoriale estendendone le caratteristiche di complementarità intrinseca, ovvero il prodotto risulta nullo se i vettori a e b sono ortogonali tra loro mentre il prodotto vettoriale risulta nullo se i vettori a e b sono paralleli. Geometricamente si può interpretare come un piano orientato, dato dalla parte vettoriale del prodotto, di intensità pari alla sua parte scalare.

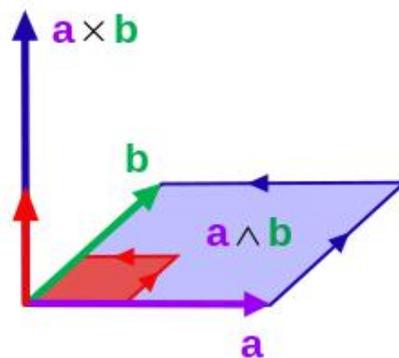


Figura 3.1: Rappresentazione geometrica del prodotto di Clifford (ab)

Grazie all'algebra di Clifford possiamo traslare operazioni geometriche in operazioni algebriche come riflessioni e rotazioni di vettori. Oltre ai nostri scopi,

una macchina basata su quest'algebra risulta uno strumento estremamente potente per applicazioni quali:

- Elaborazione dei segnali
- Geometria proiettiva
- Realtà virtuale
- Realtà aumentata
- Sistemi cognitivi
- Applicazioni di reti neurali geometriche
- Applicazioni di cinematica dei robot

L'algebra Geometrica fornisce un mezzo per manipolare multivettori (la combinazione lineare di oggetti di tipo differente) tenendo traccia di oggetti di differente grado in maniera simultanea, come si fa con le operazioni sui numeri complessi.

Algoritmo di rimozione del makeup

La figura 3.2 mostra tre immagini in diverse scale di colori. La figura 3.2-a rappresenta l'immagine originale, la figura 3.2-b è stata creata partendo dall'originale tramite una conversione YCbCr, mentre la figura 3.2-c è stata ottenuta dall'originale tramite la conversione in scala di grigi PLS.

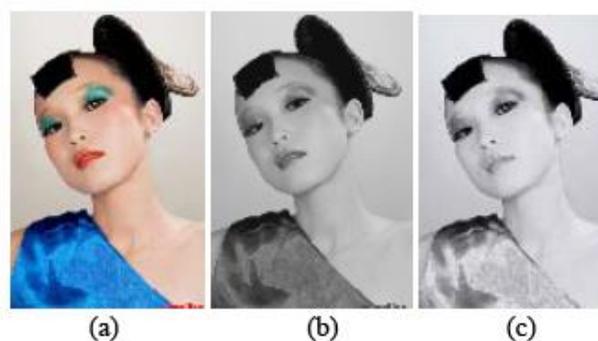


Figura 3.2: (a)immagine originale, (b) YCbCr e (c) PLS

E' evidente come le tre immagini rappresentino lo stesso individuo, con le stesse caratteristiche. L'unica differenza tra le immagini (b) e (c) riguarda l'intensità dell'immagine, differenza che l'occhio umano non può distinguere.

Il modello dell'algebra di Clifford illustrato in questo paragrafo è definito come segue:

$$Y = f_{col}(\vec{A}, \vec{b}) = y$$

f_{col} indica una funzione generale e Y , indica la somiglianza di due vettori e $b \in A^{(3,0,0)}$. La formula può essere espansa come segue:

$$f_{col}(\vec{A}, \vec{b}) = |reje_A(b)| + |k \cdot proj_A(b)|$$

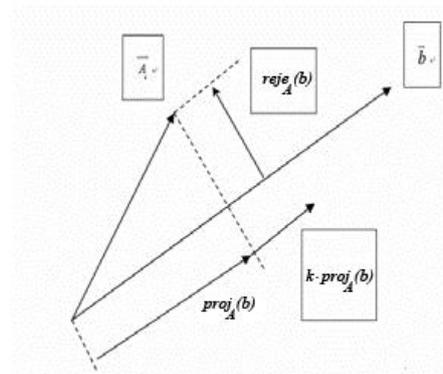


Figura 3.3: Processo di incremento dei valori della scala di grigi

In figura 3.3 si possono vedere i valori della formula riferiti al processo di incremento dei valori della scala di grigi ($reje_A(b)$ e $k \cdot proj_A(b)$).

In figura 3.4-a abbiamo l'immagine originale (\vec{A}) come quella di figura 3.2-a, in (b) l'immagine in scala di grigi (\vec{A}') come in figura 3.2-b, in figura 3.4-c è rappresentato il risultato dopo l'applicazione dell'algoritmo di rimozione del makeup basato sull'algebra di Clifford con un coefficiente di compensazione $k=0.8$. La figura 3.4-d invece è il risultato dell'applicazione del medesimo algoritmo con $k=1,5$.

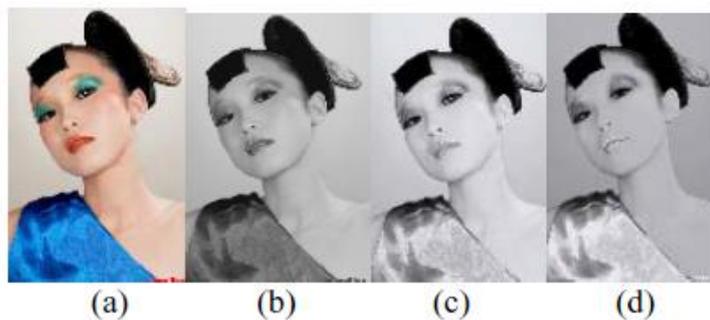


Figura 3.4: Immagini in scala di grigi dopo l'applicazione dell'algoritmo (c) e (d)

Si può vedere la differenza tra la semplice immagine in scala di grigio (figura 3.4-b) e le due immagini dopo l'applicazione dell'algoritmo (figura 3.4-c e figura 3.4-d). Si nota una notevole differenza tra le immagini (c) e (d) nelle ombre nella zona occhia a causa dei valori diversi di k .

L'efficacia dell'algoritmo può essere visto anche in figura 3.5.

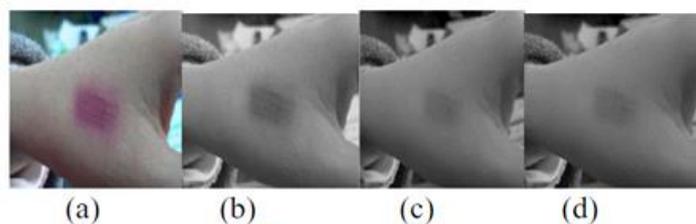


Figura 3.5: Serie di immagini. (a) immagine originale, (b) scala di grigi, (c) e (d) immagini dopo l'applicazione dell'algoritmo

La figura 3.5-a mostra l'immagine originale a colori (RGB), mentre come per figura 3.4-c e (d), le immagini 3.5-c e (d) rappresentano l'immagine (a) processata con l'algoritmo di Clifford con $k=0,8$ e $k=1,5$. E' facile notare come la mano risulta sempre più pulita dal makeup passando dalla figura (b) alla (d).

3.3 Algoritmi di confronto robusti al problema del makeup

Nel seguente paragrafo vengono analizzati tre algoritmi che effettuano un vero e proprio confronto tra due volti cercando di effettuare un riconoscimento facciale che non sia influenzato da eventuale applicazione di makeup.

3.3.1 Makeup detector applicato al riconoscimento del volto

In questo paragrafo verrà analizzata l'efficacia dell'algoritmo di rilevamento del makeup su di un volto, trattato nel paragrafo 3.1.1, applicato ad un sistema di riconoscimento facciale.

Nel paragrafo 2.4.1 si è visto quanto l'efficacia di un sistema di riconoscimento del volto possa diminuire quando si confronta un volto con makeup con uno senza. Per mitigare l'effetto negativo del trucco, l'idea è quella di inserire una fase di pre-processo nella quale si applica una normalizzazione fotometrica assieme ad un operatore di sfocatura che leviga gli effetti di contorni netti creati dal trucco.

In dettaglio quando si pongono due immagini a confronto e una delle due è identificata come avente trucco, vengono normalizzate entrambe attraverso l'utilizzo della tecnica Multiscale Self Quotient Image (MSQI) prima di essere confrontate.

L'immagine normalizzata è ottenuta secondo:

$$Q = \frac{I(x, y)}{\hat{I}(x, y)} = \frac{\rho_{\omega}(x, y)n(x, y)s}{G_k * [\rho_{\omega}(x, y)n(x, y)s]}$$

Dove Q indica l'immagine risultante, I l'immagine iniziale, $\rho_{\omega}(x,y)$ è il coefficiente di riflessione, n è la normale alla superficie, s è il riflesso dato dall'illuminazione, G_k è il filtro pesato di sfocatura Gaussiano e k è la dimensione del kernel (negli esperimenti fatti 3x3, 5x5, 11x11 e 15x15). L'immagine risultante è la somma delle immagini ottenute con i quattro valori di k .

Ancora una volta ci avvaliamo dello schema delle 5 cartelle per valutare l'efficienza dell'algoritmo con il database YMU. Per codificare e valutare l'efficienza di confronto, l'algoritmo utilizza il Multi-Scale LBP (MSLBP) [29]. MLSBP si basa su LBP ma diversamente da esso il modello binario di un pixel viene calcolato confrontando i valori medi dei sottoblocchi e a scale multiple su una griglia densa di spaziatura di 10 pixel.

A questo punto viene utilizzato l'Histogram Intersection Distance sui due vettori delle *feature* per generare un punteggio di match.

Viene riproposto lo stesso tipo di esperimento del paragrafo 2.4.1 riportando i dati in tabella 12.

Prove	M vs N	Incremento	M vs M	N vs N
1	56,25/65,55	9,3	92,86/92,86	96,43/96,43
2	52,75/55,64	2,89	73,44/80,47	87,69/87,76
3	48,54/54,00	5,46	83,33/83,33	89,23/89,29
4	45,55/49,23	3,68	80,00/80,00	92,97/95,74
5	54,34/56,35	2,01	88,46/88,46	95,73/96,15
Globale	48,88/54,10	5,22	84,70/86,05	92,72/92,72

Tabella 12: Risultati del match

Nella tabella 12 sono riportati i risultati prima e dopo l'applicazione del metodo (a/b). Come specificato la fase di pre-processo viene applicato quando una delle due immagini confrontate è etichettata come avente makeup e l'altra no. La terza colonna mostra il miglioramento apportato dopo l'applicazione della fase di pre-processo.

3.3.2 Riconoscimento facciale (LBP + DT-CWT)

Il sistema di riconoscimento facciale presentato in questo paragrafo prevede l'estrazione delle *feature* tramite l'uso di LBP e del Dual-Tree Complex Wavelet Transform (DT-CWT) [20] (figura 3.13). Successivamente le *feature* estratte vengono combinate per creare un unico vettore delle *feature* che sarà classificato da SVM. Per verificare l'efficacia del metodo vengono riportati i dati di altre famose tecniche di estrazioni di *feature*.

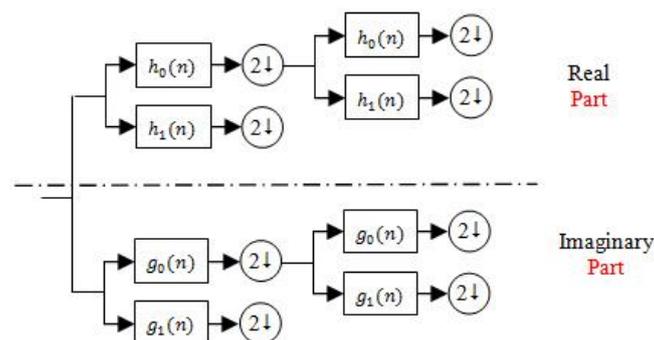


Figura 3.13: Banco di filtri 1D DT-CWT

2D DT-CWT è in grado di rilevare (in un'immagine 2D) le varie sfumature dei dettagli a sei angoli ± 15 , ± 45 , ± 75 di un volto.

Per ottenere il nostro vettore di *feature* il sistema svolge i seguenti passaggi (figura 3.14):

- Nell'immagine inserita in input, viene individuato il volto e i suoi *landmark* tramite CLM (Constrained Local Model) [32]. DT-CWT viene applicato al volto e così vengono generate 8 nuove immagini (due approssimazioni del volto e 6 con i dettagli tramite rilevazione dei bordi nei sei angoli ± 15 , ± 45 , ± 75).

- LBP viene applicato alle 8 immagini appena create.
- Viene creato un unico vettore finale.

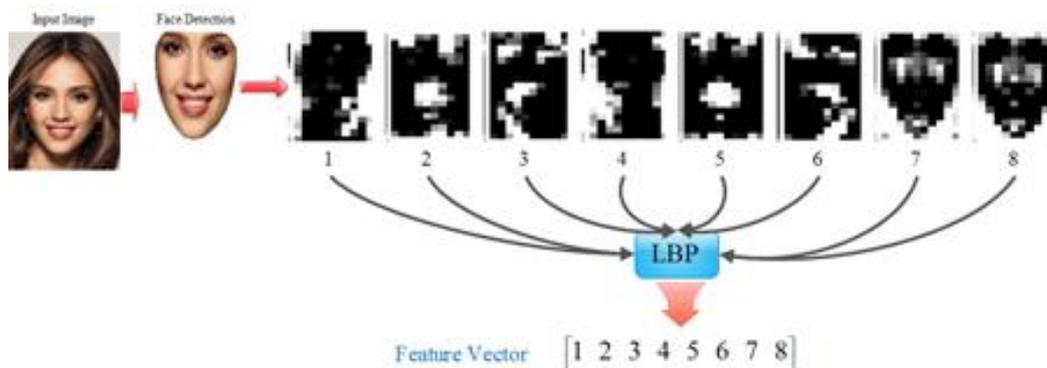


Figura 3.14: Creazione del vettore tramite DT-CWT + LBP

Il sistema di riconoscimento facciale di basa su due processi, uno offline e uno online. Nel processo offline vengono generati (secondo il procedimento appena descritto) dei vettori per ogni volto senza makeup in possesso. A questo punto, nella fase di training, vengono generate tante classi quanti sono i soggetti analizzati.

Durante la fase online viene creato (sempre seguendo lo schema di figura 3.14) il vettore dall'immagine con makeup di cui si vuole effettuare il riconoscimento.

A questo punto il vettore viene classificato tramite l'utilizzo di SVM confrontandolo con le classi generate offline (figura 3.15).

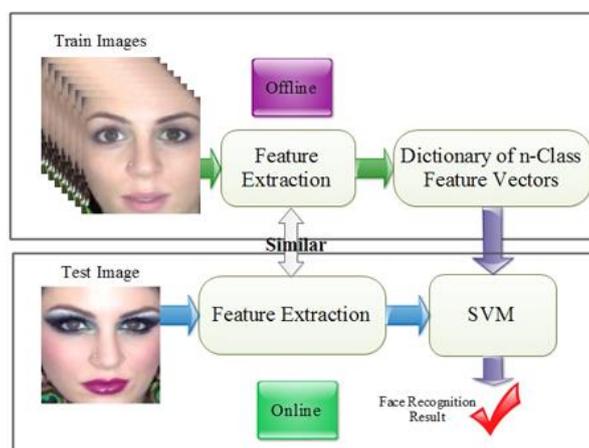


Figura 3.15: Processo di riconoscimento di un volto

Esperimento 1 su YMU

Di seguito sono riportati i risultati dell'algorithmo negli esperimenti condotti da Moeini H., Mozzaffari e Moeini A. [24] confrontando il metodo con altre tecniche di estrazione di feature.

Tecnica	% di riconoscimento
DT-CWT + LBP	91,6%
DT-CWT	84,1%
Gabor	86,7%
LGBP	80,5%
LBP	74,4%
PCA	71,0%
HOG	71,9%

Tabella 13: Risultati degli esperimenti su YMU

Esperimento 2 su VMU

Data la struttura del database VMU (paragrafo 2.3.2) per il secondo esperimento [24] sono stati fatti tre tipi di test: immagini con solo il rossetto (a), trucco solo nella zona occhi (b) e trucco completo (c).

Tecnica	% (a)	% (b)	% (c)	Media
DT-CWT + LBP	92,0%	94,0%	90,0%	92,0%
DT-CWT	86,2%	88,2%	84,2%	86,2%
Gabor	86,2%	84,2%	84,2%	84,8%
LGBP	84,2%	82,2%	80,2%	82,2%
LBP	80,2%	78,3%	76,3%	78,3%
PCA	78,3%	78,3%	74,3%	76,9%
HOG	74,3%	74,3%	72,3%	73,6%

Tabella 14: Risultati degli esperimenti su VMU

Conclusioni

In entrambi gli esperimenti condotti, la combinazione di DT-CWT e LBP risulta essere molto più efficiente rispetto ad altri metodi con approcci simili, perfino in ogni tipologia di test svolta nell'esperimento 2 (a, b, c).

3.4 VeriLook 5.3/MegaMatcher 4.4 Algorithm

In questo paragrafo verrà usato un software commerciale, nella sua versione demo, per testarne l'efficacia utilizzando come dataset YMU e VMU.

In particolare, per gli esperimenti effettuati in questo paragrafo, il software verrà lanciato su di un sistema Windows 7 con 4 GB di RAM e 2.30 GHz di CPU.

VeriLook 5.3/MegaMatcher 4.4 Algorithm Demo è un'applicazione progettata con l'obiettivo di mettere alla prova le potenzialità del motore di riconoscimento facciale Neurotechnology. Questo software è in grado di utilizzare immagini acquisite da sorgenti video esterne (come una camera IP), oppure rilevare volti direttamente da immagini (supportando varie estensioni come .bmp, .tif, .png, .jpg o .gif).

L'applicazione permette 3 operazioni fondamentali:

- **Enrollment:** durante questa fase, il software elabora il volto dell'immagine estraendone le feature e memorizzandole in un database per successivi confronti.
- **Face enrollment feature generalization:** questa modalità genera delle feature del volto generalizzate partendo da un numero di template della stessa persona. Ogni immagine di volto viene processata e le sue feature estratte. Successivamente vengono analizzate delle collezioni di feature e combinate in una collezione di feature generalizzata scritta nel database.

La qualità del riconoscimento facciale incrementa utilizzando questa modalità.

- **Matching:** attraverso questa modalità è possibile confrontare una nuova immagine con quelle memorizzate nel database attraverso l'operazione di enrollment.

3.4.1 Requisiti di qualità delle immagini

L'accuratezza del sistema di riconoscimento dipende in grande misura dalla qualità dell'immagine del volto. In modo particolare bisogna prestare maggior attenzione alla qualità durante la fase di enrollment, dove una o più immagini vengono convertite in template di feature del volto per poi essere salvate nel database per essere così confrontate in maniera rapida con un'immagine da identificare.

Gli sviluppatori del software suggeriscono in particolare modo di prestare principalmente attenzione a tre requisiti:

- **Posizione del volto ed espressioni:** per maggiori prestazioni è consigliabile usare immagini frontali ed in primo piano di un volto ruotato fino ad un massimo di 15 gradi in ogni direzione. Usare più immagini durante la fase di enrollment per colmare le differenze causate dai vari angoli di inclinazione di un volto. Durante la fase di enrollment sono preferite espressioni facciali neutre, mentre durante la fase di identificazione tali espressioni non inficiano significativamente la precisione dei risultati. Come già accennato, l'uso di occhiali da sole diminuisce drasticamente la qualità dei risultati, mentre per garantire l'efficacia del sistema anche in presenza di occhiali da vista è consigliabile inserire, durante la fase di enrollment, immagini sia con gli occhiali sia senza. In ultimo, la distanza minima tra gli occhi

nell'immagine dovrebbe essere di 50 pixel, mentre il valore raccomandato è di 75 pixel.

- **Luce:** nel caso in cui le condizioni di luminosità possano essere controllate, la luce dovrebbe essere distribuita equamente su ogni lato del volto e anche dalla parte bassa a quella alta del viso, senza significative zone d'ombra. Questi requisiti possono essere raggiunti attraverso l'uso di una luce frontale diretta o con l'uso di una luce diffusa. Quindi è preferibile evitare la luce del sole o luce che può produrre riflessi sulle lenti degli occhiali o sulla pelle.
- **Camera:** per l'acquisizione delle immagini dei volti è consigliato l'uso di dispositivi che producano risultati dalla qualità simile sia per la fase di enrollment, sia per quella di identificazione. Assicurarsi che i volti vengono acquisiti nello stesso modo: alcuni dispositivi di acquisizione possono essere configurati per produrre immagini specchiate. Accertarsi che tutte le immagini siano acquisite con lo stesso metodo, perché i template delle feature estratte da un'immagine specchiata da una medesima persona potrebbero non combaciare con quelle estratte da un'immagine non specchiata.

Negli esperimenti che eseguiremo per testare il software non dovremo porci il problema di creare un ambiente controllato, in quanto utilizzeremo dei dataset già testati con altri algoritmi, così da poter fare un confronto.

3.4.2 Configurazione del software

In questo paragrafo verranno impostate le opzioni del software per raggiungere il miglior risultato. Sono disponibili quattro schede di configurazione:

- Face detection: contiene le opzioni per il rilevamento di un volto nell'immagine.
- Face extraction: riguarda il processo di estrazione delle feature da un volto.
- Enrollment: opzioni che riguardano la fase di enrollment.
- Identification: opzioni riguardanti la fase di riconoscimento di un volto cercando tra quelli caricati nel database.

Face detection

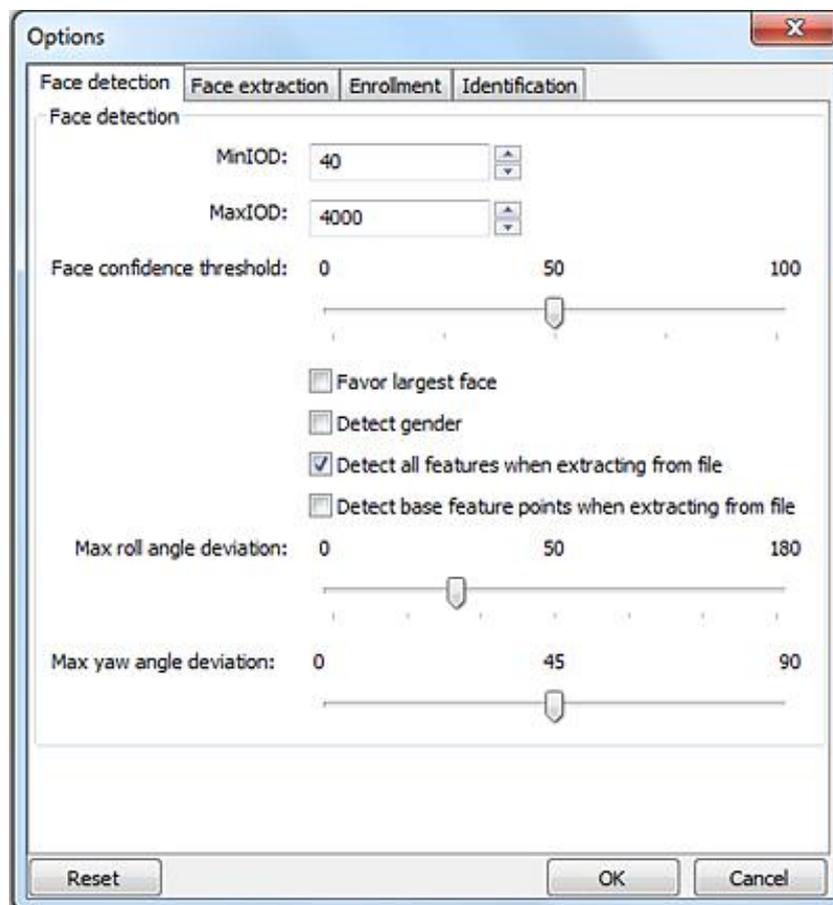


Figura 3.16: impostazioni del Face detection

La scheda è stata impostata come in figura 3.16. Di seguito vengono elencate le impostazioni a disposizione.

Minimum IOD / Maximum IOD: è la minima/massima distanza tra gli occhi ed è stata lasciato come valore quello di default, 40 / 4000.

Face confidence threshold: il valore indica i requisiti che deve possedere un volto per essere riconosciuto durante la fase di rilevamento. Maggiore è il valore e più restrittive saranno le regole applicate durante la ricerca.

Favor largest face: se abilitata, l'opzione cerca nell'immagine il volto più grande e sceglie quello per estrarne le feature piuttosto che quello che ottiene il punteggio più alto. Nel nostro caso abbiamo delle immagini contenenti un unico volto in primo piano quindi la spunta o meno non cambia il risultato.

Detect gender: cerca di riconoscere anche il sesso dell'individuo.

Detect all feature when extracting from file: vengono rilevate tutte tutte le feature da un volto e non solo quelle strettamente necessarie per l'estrazione.

Detect base feature points when extracting from file: rileva occhi, punta del naso e bocca.

Max roll angle deviation / Max yaw angle deviation: indica il massimo valore in gradi in cui un volto può essere ruotato per essere processato (sia durante la fase di enrollment sia durante quella di identificazione). Rispetto a quelle di default i valori sono stati aumentati per acquisire più volti (50 e 45).

Face extraction

In questa scheda si può impostare, in una scala da 0 a 255, il valore di qualità dell'immagine richiesta per far sì che essa venga processata e quindi ne vengano estratte le feature. Siccome le nostre immagini dei volti sono piccole e di scarsa qualità, per fare in modo che il software le processi è necessario impostare il valore a 0, altrimenti verrà restituito per la maggior parte delle immagini un errore di quality exposure check failed.

Enrollment

Questa scheda riguarda la fase di estrazione delle feature durante la fase di enrollment. Le opzioni di questa scheda sono state lasciate quelle di default (figura 3.17), in particolare sono state lasciate le due spunte e le dimensioni del template estratto da un'immagine è stato lasciato Large. Il resto delle

impostazioni non importano in quanto non verranno utilizzate per i nostri esperimenti.

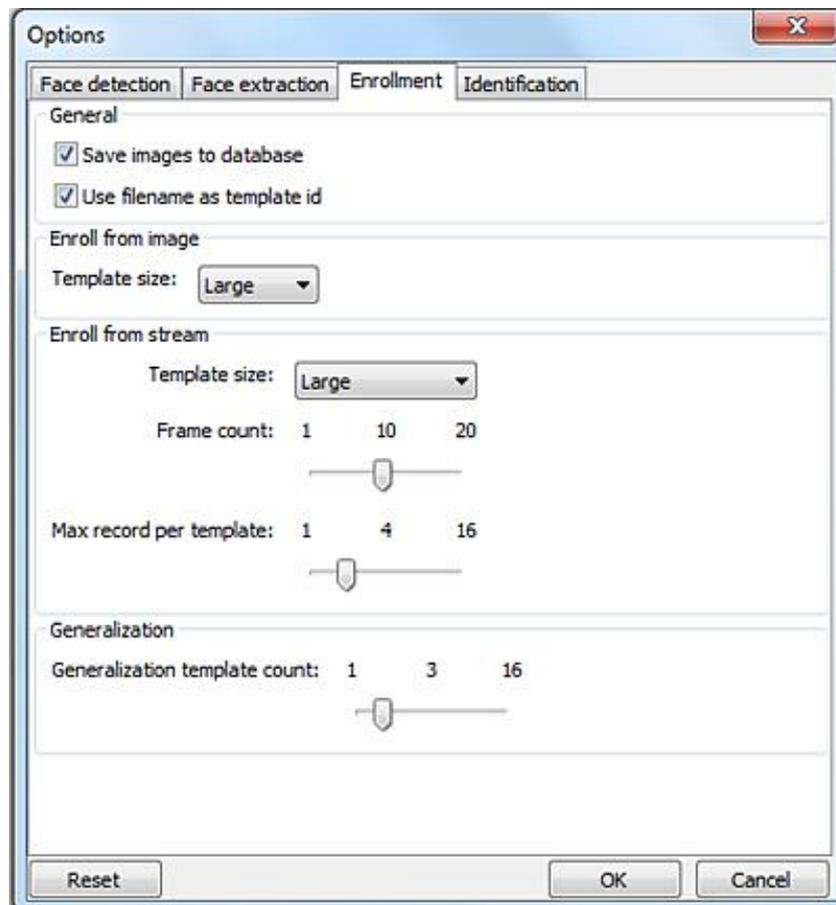


Figura 3.17: impostazioni della fase di Enrollment

Identification

Le opzioni della scheda identification sono state impostate come in figura 3.18. In particolare modo, come suggerito dagli sviluppatori, le dimensioni del template estratto durante la fase di identificazione è stato lasciato a Medium. Le opzioni di Identify from stream sono state lasciate inalterate in quanto i volti necessari per gli esperimenti verranno acquisiti da file e non da stream.

Infine due fattori importanti per la fase di identification sono la velocità di matching e l'opzione FAR (false acceptance rate). Per consentire un match più accurato è consigliabile impostare la velocità a Low, mentre tramite FAR impostiamo una soglia sotto la quale il risultato non viene preso in esame. Il

punteggio di match è strettamente legato alla soglia FAR. Più il valore di match è alto e più il valore FAR sarà basso. Di seguito viene riportata una tabella di conversione FAR-score (tabella 15).

FAR (false acceptance rate)	score
1%	24
0.1%	36
0.01%	48
0.001%	60
0.0001%	72
0.00001%	84
0.000001%	96

Tabella 15: tabella di conversione FAR-score

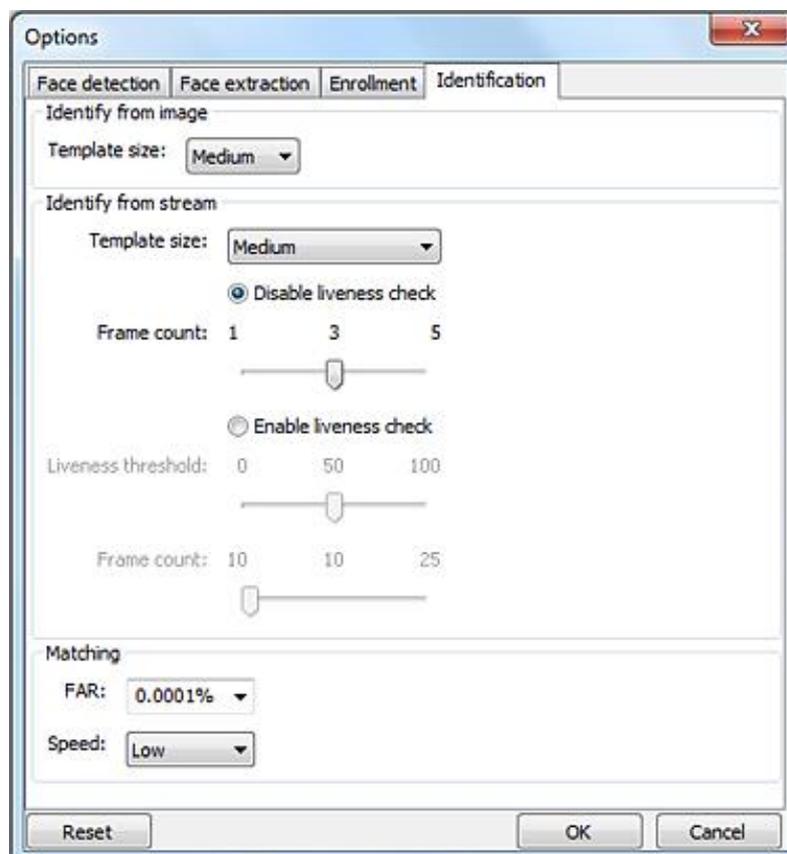


Figura 3.18: impostazioni della fase di Identification

I valori sono ottenuti tramite la formula:

$$Score = -12 * \log_{10} \left(\frac{FAR}{100} \right)$$

Nell'ambito dei documenti di identità il valore FAR viene solitamente impostato a 0,1%. Nel nostro esperimento viene usato come valore FAR 0.0001%, ciò significa che tutte le immagini del database che non raggiungono un punteggio di almeno 72 non vengono mostrate a video.

3.4.3 Esperimenti

Di seguito verranno riportati i risultati degli esperimenti divisi in base al dataset di immagini utilizzato, effettuando per YMU prima la fase di enroll sulle immagini prive di makeup procedendo poi alla fase di identificazione usando le immagini con makeup e poi effettuando l'enroll sulle immagini con makeup e l'identificazione attraverso quelle prive di makeup, proponendo i relativi risultati. Mentre per VMU vengono effettuati dei confronti tra volti senza makeup con volti con diversi tipi di trucco per valutare quale cosmetico e in che misura deteriora maggiormente l'efficacia del riconoscimento.

Esperimento su YMU

Sono state prese le 302 immagini del dataset prive di makeup e tramite la funzione enroll, le feature sono state caricate nel database del software. Di queste 302 immagini caricate però, solamente 281 sono state effettivamente rilevate come volti ed inserite nel database. A causa della qualità, del volto non frontale o non perfettamente centrato, 21 immagini sono state scartate (un esempio in figura 3.19).

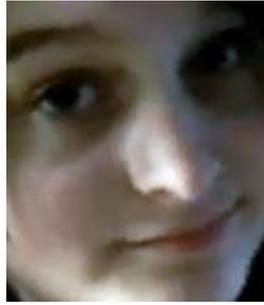


Figura 3.19: volto non riconosciuto dal software

A questo punto sono state inserite una alla volta le immagini dei volti con il makeup con la funzione identify ed è stato preso nota del risultato.

Delle immagini con il makeup inserite nel software, di 27 non è stato individuato il volto, di 14 non è stata riscontrata nessuna corrispondenza con le 281 immagini senza makeup precedentemente caricate nel database del software, mentre le restanti immagini sono state processate e per ognuna è stato restituito in output una classifica in ordine di similarità. In questo esperimento si è preso in considerazione in che posizione della classifica la vera immagine senza makeup, corrispondente a quella inserita, si trova. In figura 3.20 si riporta il grafico rappresentante la capacità del software di riconoscere con precisione il giusto individuo tra quelli del database caricato. In ascisse si trovano le posizioni della classifica creata dal software, mentre in ordinata la percentuale di immagini trovate nella relativa posizione.

Come viene mostrato anche nel grafico, il software ha riconosciuto esattamente nella prima posizione il giusto individuo 151 volte delle 282 processate (ovvero ha avuto successo il 53.6% delle volte). Il resto dei risultati si distribuisce secondo l'andamento visibile dal grafico. Nella tabella 16 vengono mostrate le prime venti posizioni, che sono le posizioni con la più alta concentrazione di ricorrenze. Dopo le prime venti posizioni le occorrenze tendono a diradarsi molto. A differenza del grafico 3.20, in tabella 16 sono state inserite le effettive occorrenze per ogni posizione invece del valore cumulativo.

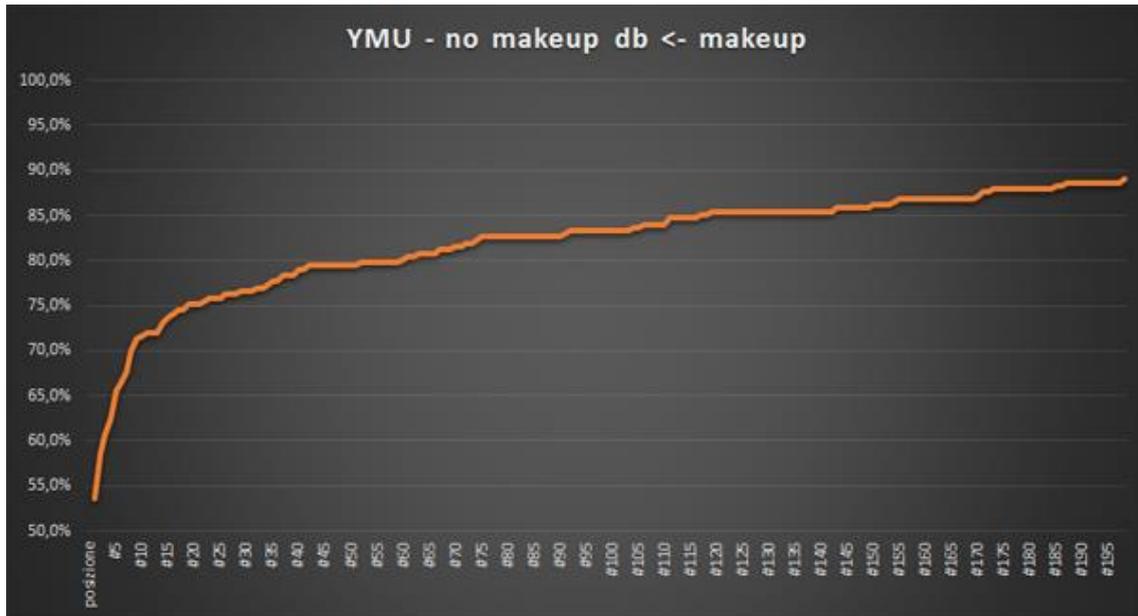


Figura 3.20: grafico dei risultati dell'esperimento su YMU (prima parte)

Posizione	#Occorrenze	Posizione	#Occorrenze
#1	151	#11	1
#2	14	#12	0
#3	6	#13	0
#4	5	#14	3
#5	9	#15	2
#6	2	#16	1
#7	4	#17	1
#8	6	#18	0
#9	4	#19	2
#10	1	#20	0

Tabella 16: prime 20 posizioni con relative occorrenze della prima parte dell'esperimento

Come preannunciato è stato poi eseguito l'esperimento invertendo i passaggi: dopo aver azzerato il database del software, sono state caricate, attraverso la funzione di enroll, le immagini dei volti con il makeup e successivamente sono

state confrontate manualmente, un'immagine alla volta, le immagini senza makeup. A differenza dell'esperimento precedente, sono state confrontate un numero inferiore di immagini, perché il dato interessante che si voleva verificare era se la distribuzione delle occorrenze seguiva lo stesso del precedente, ovvero se la scelta di applicare l'enroll ai volti con makeup o a quelli senza influenzava i risultati.

Durante questa seconda fase sono state confrontate con il database precaricato 160 immagini senza makeup: di queste in 12 non è stato riscontrato un volto, in 7 non è stato trovato l'effettivo volto corrispondente col makeup, mentre 82 sono state correttamente identificate al primo posto (ottenendo una percentuale di successo del 51.3%, molto vicina a quella precedente). In tutti e due i test si vede chiaramente che la curva non raggiunge mai il 100% delle immagini, manca infatti circa un 10% dovuto alle immagini non riconosciute o non processate. Questo porta alla conclusione che qualsiasi sia l'ordine in cui si sceglie di procedere durante l'esperimento il risultato non varia. Come prima, di seguito vengono proposti grafico (figura 3.21) e tabella delle prime 20 posizioni (tabella 17) per cogliere, anche a colpo d'occhio, i risultati raggiunti.

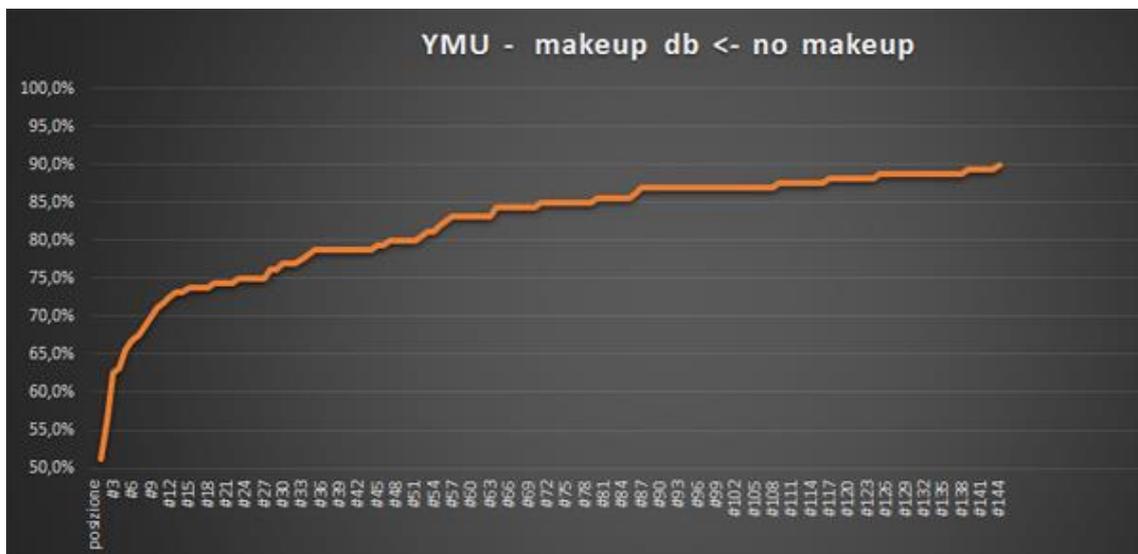


Figura 3.21: grafico dei risultati dell'esperimento su YMU (seconda parte)

Posizione	#Occorrenze	Posizione	#Occorrenze
#1	82	#11	1
#2	8	#12	1
#3	10	#13	1
#4	1	#14	0
#5	4	#15	1
#6	2	#16	0
#7	1	#17	0
#8	2	#18	0
#9	2	#19	1
#10	2	#20	0

Tabella 17: prime 20 posizioni con relative occorrenze della seconda parte dell'esperimento

Esperimento su VMU

Il secondo esperimento è stato condotto utilizzando VMU come dataset. A differenza del primo esperimento, utilizzando VMU, non si analizzerà l'efficienza del software utilizzato, ma testeremo con quale tipo di trucco il software riscontra più difficoltà, per confrontare poi i risultati con l'esperimento di Dantcheva (paragrafo 2.4.2). Mentre i vari volti con makeup (labbra, occhi e makeup totale) sono ottenuti tutti dalla stessa immagine, garantendo così pari condizioni di luminosità e posizione per i vari tipi di trucco, i volti senza makeup sono ottenuti in pose leggermente diverse. Per prima cosa è stato effettuato l'enrollment delle 51 immagini senza trucco in posa diversa rispetto a quelle col trucco. Poi sono stati eseguiti i confronti con le 51 immagini sempre senza trucco ma nella stessa posa delle immagini a cui poi è stato aggiunto il trucco. Durante questa prima fase il software si è comportato abbastanza bene, riconoscendo correttamente 46 individui su 51 e a parte un unico caso (in cui il soggetto corretto era in posizione 30) le restanti 4 immagini sono state collocate vicino alle prime posizioni (riepilogo dei risultati nella tabella 18).

Posizione	#Occorrenze
#1	46
#3	1
#5	1
#6	1
#11	1
#30	1

Tabella 18: risultato dell'esperimento su VMU no-makeup con no-makeup

A questo punto si procede verificando se l'aggiunta di trucco alle labbra peggiori l'efficienza del software o meno. Quindi vengono confrontate una ad una le immagini con il rossetto al database già inserito nel software. I risultati di questo esperimento mostrano un lieve calo di efficienza del programma, riuscendo ad individuare correttamente l'individuo 44 volte su 51 (tabella 19), rispetto alle 46 in assenza di trucco. In aggiunta per un'immagine, dopo l'aggiunta del rossetto, il sistema non è stato più in grado di individuare il volto.

Posizione	#Occorrenze
#1	44
#2	2
#7	2
#19	1
#28	1
Volto non trovato	1

Tabella 19: risultato dell'esperimento su VMU no-makeup con rossetto

A questo punto viene testato quanto influenzi un sistema di riconoscimento del volto il trucco nella sola zona occhi. Come ci si aspettava la, zona occhi diminuisce la precisione del software in maniera maggiore rispetto alle labbra. Infatti il test ha successo solamente 42 volte su 51, inoltre 2 volte non viene

rilevato nessun volto e 1 volta il soggetto corretto non viene nemmeno messo in classifica (risultati in tabella 20).

Posizione	#Occorrenze
#1	42
#4	2
#6	1
#7	1
#21	1
#23	1
Nessun match	1
Volto non trovato	2

Tabella 20: risultato dell'esperimento su VMU no-makeup con trucco occhi

Posizione	#Occorrenze
#1	39
#2	2
#3	1
#4	1
#7	1
#16	1
#24	1
Nessun match	2
Volto non trovato	3

Tabella 21: risultato dell'esperimento su VMU no-makeup con trucco completo

Infine si è ripetuto il test confrontando il database con le immagini alle quali è stato aggiunto un makeup totale. Anche in questo caso abbiamo un ulteriore

peggiore in termini di efficienza del software. Sono stati individuati correttamente 39 soggetti su 51, non sono stati individuati ben 3 volte e per 2 volte il soggetto analizzato non è stato inserito in classifica dal software (tabella 21).

Conclusioni

Dal primo esperimento sul dataset YMU si deduce che il software in questione è molto sensibile alla qualità dell'immagine analizzata e alla variazione di luminosità e posa. Con un risultato di circa 54.9% di successo, il software non risulta abbastanza attendibile da poter essere applicato in settori che riguardano la sicurezza. Nel secondo esperimento, dove posa e illuminazione si possono considerare irrilevanti, riusciamo a decretare quale tipo di trucco genera più errori di valutazione e in quale misura. I risultati confermano i dati ottenuti da Dantcheva nel suo esperimento (paragrafo 2.4.2), ovvero il rossetto non è un cosmetico che altera drasticamente le caratteristiche di un volto. Mentre il trucco agli occhi e ovviamente quello totale risultano essere fonte di errori più gravi per il sistema di valutazione del software.

Conclusioni

In questo lavoro di tesi si è studiato l'effetto che l'applicazione di trucchi cosmetici può avere su di un sistema di riconoscimento del volto. Questo tema risulta molto importante per il settore della sicurezza, in quanto l'uso di tecniche di makeup, a differenza di altre tecniche con fini criminali, richiede strumenti alla portata di tutti ed è un metodo molto semplice per cercare di ingannare dei sistemi di riconoscimento biometrici basati sulle caratteristiche principali di un viso.

In particolare si sono analizzate le tecniche più diffuse di riconoscimento del volto senza considerare il problema del makeup, focalizzandosi sul loro funzionamento e sulla loro efficienza.

Successivamente ci si è focalizzati sull'uso dei cosmetici analizzando come un makeup possa essere applicato (sia reale che virtuale) e sui risultati di studi che provano quanto l'applicazione di semplice trucco possa alterare l'efficienza di un sistema di riconoscimento.

Infine si sono analizzati gli algoritmi maggiormente noti in letteratura. Questi algoritmi hanno diversi obiettivi: individuare la presenza del trucco, cercare di eliminarlo digitalmente ed infine eseguire dei veri e propri confronti tra individui con e senza makeup. Oltre all'analisi di questi algoritmi, sono stati condotti degli esperimenti per verificare se il software commerciale VeriLook 5.3/MegaMatcher 4.4 Algorithm Demo possa essere considerato un valido mezzo in grado di gestire i cambiamenti apportati al volto dai cosmetici.

Un unico software si rivela non essere una soluzione a questo problema in quanto restituisce dei risultati non accettabili in un campo delicato come quello della sicurezza pubblica. Mentre analizzando i vari algoritmi si evince che la soluzione migliore risulta essere quella di combinare più algoritmi assieme. Nonostante in questo modo i risultati migliorino drasticamente, non sono così buoni da poter

affermare che essi siano sistemi sicuri contro dei tentativi di mistificazione dei tratti di un volto a scopi criminali.

In definitiva il problema dei sistemi di riconoscimento del volto in presenza di makeup risulta essere ancora un problema aperto, ed è certamente necessario lo studio di nuovi algoritmi che raggiungano un migliore livello di robustezza e accuratezza.

Ringraziamenti

Desidero ringraziare la Dott.ssa Annalisa Franco per la sua disponibilità e cortesia dimostratami durante la stesura di questa tesi e, in generale, durante tutto il periodo di studi.

Un ringraziamento particolare va ai miei genitori, i quali mi hanno sostenuto moralmente e incoraggiato a non mollare mai, anche nei periodi di maggiore difficoltà, dimostrandomi un immenso affetto.

Ringrazio anche la mia morosa Giorgia, sempre al mio fianco, capace di esaltare le gioie dei miei successi e di minimizzare la tristezza delle delusioni.

Desidero ringraziare anche i miei amici, Marco e Massimo, con i quali ho condiviso una parte fondamentale del percorso di studi e grazie ai quali ho portato a termine dei progetti di cui vado tutt'ora fiero.

Infine ringrazio tutti quelli che hanno sempre creduto in me e che fanno parte della mia vita.

Bibliografia e Sitografia

[1] Adkins L. e Adkins R., “*Handbook to Life in Ancient Greece*”. Oxford, U.K.: Oxford Univ. Press, 1998.

[2] Angelina Jolie Makeup Tutorial
<https://www.youtube.com/watch?v=rNQ2oDEyVNw>

[3] Angeloglou M., “*A History of Make-up*”. Macmillan, 1970

[4] Ballantyne M., R. S. Boyer e L. Hines (1996). “*Woody bledsoe: His life and legacy.*” AI Magazine.

[5] Biglari M., F. Mirzaei, e H. Ebrahimpour-Komeh, “*Illumination invariant face recognition using SQI and weighted LBP histogram*” in 2013 First Iranian Conference on Pattern Recognition and Image Analysis (PRIA), 2013.

[6] Brand J., Mason J.S. (2000) “*A Comparative Assessment of Three Approaches to Pixel-level Human Skin-Detection*”

[7] Chang C.-C. e Lin C.-J. “*LIBSVM: A library for support vector machines*”, ACM Trans. on Intelligent Systems and Technology, vol. 2, 2011

[8] Chen C., Dantcheva A. e Ross A. ”*Automatic Facial Makeup Detection with Application in Face Recognition*”, Int. Conf. on Biometrics (ICB), 2013.

[9] Dantcheva A. e Dugelay J., “*Female facial aesthetics based on soft biometrics and photo-quality*” in Proc. IEEE Conf. Multimedia and Expo, 2011.

[10] Dantcheva A., Chen C. e Ross A. “*Can facial cosmetics affect the matching accuracy of face recognition systems?*” In BTAS, 2012.

[11] Eckert M.-L., Kose N. e Dugelay J.-L. “*Facial cosmetics database and impact analysis on automatic face recognition*”, IEEE Int. Workshop on Multimedia Signal Processing (MMSP), 2013.

[12] Facial Cosmetics Database <http://fcd.eurecom.fr/>

[13] Federal Bureau of Investigation, https://www.fbi.gov/about-us/cjis/fingerprints_biometrics/biometric-center-of-excellence/files/face-recognition.pdf

- [14] Gu'eguen N., "The effects of women's cosmetics on men's courtship behavior" North Amer. J. Psychol., vol. 10, no. 1, 2008
- [15] Guo G., L. Wen e S. Yan, "Face Authentication With Makeup Changes" IEEE Trans. Circuits Syst. Video Technol., vol. 24, May 2014.
- [16] Hu J., Y. Ge, et al. "Makeup-robust face verification", ICASSP, 2013
- [17] INFace
http://luks.fe.uni-lj.si/sl/osebje/vitomir/face_tools/INFace/index.html
- [18] Kose N., Apville L. e Dugelay J. "Facial makeup detection technique base on texture and shape analysis" 11th IEEE International Conference on Automatic Face and Gesture Recognition, Maggio 2015
- [19] Liu C. e Wechsler H. "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition". IEEE Trans. on Image Processing, 2002.
- [20] Liu Chao-Chun, Dao-Qing Dai, "Face Recognition Using Dual-Tree Complex Wavelet Features" Image Processing, IEEE Transactions on, Nov. 2009.
- [21] Ludwig O., D. Delgado, et al. "Trainable Classifier-Fusion Schemes: An Application To Pedestrian Detection", Int. IEEE Conf. On Intelligent Transportation Systems, 2009.
- [22] Mary Kay Virtual Makeover, http://www.marykay.com/en-US/TipsAndTrends/MakeoverAndBeautyTools/_layouts/MaryKayCoreTipsAndTrends/VirtualMakeOver.aspx
- [23] Min R., Hadid A. e Dugelay J.-L., "Improving the recognition of faces occluded by facial accessories" in FG 2011
<http://www.eurecom.fr/en/publication/3314/download/mm-publi-3314.pdf>
- [24] Moeini H., Mozzaffari S. e Moeini A. "Makeup-Invariant Face Recognition by combination of Local Binary Pattern and Dual-Tree Complex Wavelet Transform from Women's Images", International Symposium on Telecommunications, 2014.
- [25] NPD https://www.npd.com/wps/portal/npd/us/news/press-releases/pr_120301/
- [26] Ogihara A., Shintani A., Takamatsu S., Igawa S. (1996), "Speech Recognition based on the Fusion of Visual and Auditory Information using Full-Frame Colour image", IEICE Trans. Fundamentals

- [27] Oliva A. and A. Torralba. “*Modeling the shape of the scene: A holistic representation of the spatial envelope*”. IJCV, 2001.
- [28] Organizzazione Internazionale dell’Aviazione Civile,
<http://www.icao.int/Security/mrtd/Pages/24-NOV-2015.aspx>
- [29] Paris S., H. Glotin, e Z. Zhao. “*Real-time face detection using integral histogram of multi-scale local binary patterns*”. In ICIC, 2011.
- [30] Rasti S., M.a. Masnadi-Shirazi e G. Elahi. “*Makeup Detection Using Local Fisher Discriminant Analysis*”. International Conference on Information and Knowledge Technology, 2015.
- [31] Rhodes G., Sumich A. e Byatt G., “*Are Average Facial Configurations Attractive Only Because of Their Symmetry?*” Psychological Science, vol. 10, 1999.
- [32] Saragih J., S. Lucey, J. Cohn, “*Deformable Model Fitting by Regularized Landmark Mean-Shift*” International Journal of Computer Vision, vol. 91, 2011.
- [33] Sirovich L. e Kirby M. (1987), “*A Low-Dimensional Procedure for the Characterization of Human Faces*” J. Optical Soc. Am. A.
- [34] Tan X. e Triggs B. “*Enhanced local texture feature sets for face recognition under difficult lighting conditions*”. IEEE Trans. on Image Processing, 2010.
- [35] Ueda S. e Koyama T. “*Influence of make-up on facial recognition. Perception*”, 2010.
- [36] Welling M., “*Fisher Linear Discriminant Analysis*”, Department of Computer Science University of Toronto
- [37] Verilook Face Toolkit <http://www.neurotechnology.com/verilook.html>
- [38] [Wikipedia,](https://it.wikipedia.org/wiki/Organizzazione_internazionale_dell%27aviazione_civile)
https://it.wikipedia.org/wiki/Organizzazione_internazionale_dell%27aviazione_civile
- [39] Wikipedia,
https://it.wikipedia.org/wiki/Sistema_di_riconoscimento_biometrico
- [40] Wikipedia, https://en.wikipedia.org/wiki/Mary_Kay
- [41] Wikipedia,
http://en.wikipedia.org/wiki/Viola%E2%80%93Jones_object_detection_framework

[42] Wikipedia,
https://it.wikipedia.org/wiki/Analisi_delle_componenti_principali

[43] Wikipedia, https://en.wikipedia.org/wiki/Canonical_correlation

[44] Zhang W., Shan S., et al. “*Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition*”, Proc. ICCV, 2005