

# Progetto 3

## Computer Vision 2020

Davide Nigri IN2000105

<b>Descrizione del problema</b>	<b>2</b>
<b>Descrizione dell'approccio</b>	<b>2</b>
<b>Scelte di implementazione</b>	<b>2</b>
<b>Risultati</b>	<b>3</b>
Fase 1	3
Fase 2	7
Data augmentation	7
Batch normalization	9
Incremento del numero e della dimensione dei filtri	11
Incremento della dimensione del minibatch	13
Incremento del numero di strati convoluzionali e dropout	13
Fase 3	14
<b>Riferimenti</b>	<b>15</b>

# Descrizione del problema

È stata richiesta la realizzazione di un classificatore di immagini mediante rete neurale convoluzionale che, data un'immagine, fosse in grado di stabilire la sua appartenenza ad una tra le seguenti quindici possibili categorie: *office, kitchen, living room, bedroom, store, industrial, tall building, inside city, street, highway, coast, open country, mountain, forest, suburb*. È stato fornito un dataset contenente 4485 immagini a livelli di grigio già etichettate e suddivise in 1500 immagini da utilizzare durante l'addestramento e 2985 durante la fase di test.

## Descrizione dell'approccio

Inizialmente è stata addestrata una rete neurale convoluzionale semplice con lo scopo di ottenere una soluzione di base al problema. Poi, questa soluzione è stata progressivamente migliorata fino al raggiungimento di un'accuratezza sui dati di test pari a circa il 60%. Infine, sono stati adottati due approcci di transfer learning (fine-tuning e feature extraction) che hanno permesso il raggiungimento di un'accuratezza superiore all'80%.

## Scelte di implementazione

### Fase 1

Escluse le impostazioni richieste dalla traccia, in questa fase sono stati inizialmente utilizzati tutti i valori di default, eccezion fatta per i parametri *ValidationPatience* e *InitialLearnRate*. Il primo è stato impostato a 5 con lo scopo di avere un criterio di early stopping durante l'addestramento. Il secondo, partendo dal suo valore di default di 0.01, è stato progressivamente diminuito fino a 0.0005 per consentire il raggiungimento dell'obiettivo previsto per questa fase, ovvero un'accuratezza di circa il 30%.

### Fase 2

Per incrementare le prestazioni del classificatore sono stati implementati, uno dopo l'altro, tutti i suggerimenti proposti fino a che non è stata raggiunta un'accuratezza sui dati di test di circa il 60%. In particolare, le migliori implementate sono quelle elencate di seguito. Se non diversamente specificato, i valori di eventuali parametri sono quelli di default.

- Data augmentation (left-to-right reflection)
- Batch normalization e incremento di *InitialLearnRate* a 0.001
- Incremento del numero e dimensione dei filtri nel seguente modo:

Strato convoluzionale	Prima	Dopo
conv_1	8 filtri 3x3	16 filtri 3x3
conv_2	16 filtri 3x3	32 filtri 5x5
conv_3	32 filtri 3x3	64 filtri 7x7

- Incremento della dimensione del minibatch a 64 e successivamente a 128
- Aggiunta di due strati convoluzionali, uno strato di dropout, incremento di *InitialLearnRate* a 0.005 e di *MaxEpochs* a 50

Strato convoluzionale aggiuntivo	Configurazione
conv_4	64 filtri 9x9
conv_5	64 filtri 11x11

### Fase 3

Sia per il fine-tuning che per la procedura di feature extraction è stata utilizzata la rete pre-addestrata AlexNet.

Nel primo caso, ai parametri *InitialLearnRate* e *MaxEpochs* sono stati assegnati rispettivamente i nuovi valori 0.001 e 15. Per forzare la rete ad addestrare principalmente l'ultimo strato completamente connesso, sono stati impostati a 50 i parametri *WeightLearnRateFactor* e *BiasLearnRateFactor*.

Nel secondo caso, la classificazione finale è stata portata a termine sia con un modello Pairwise SVM che con un modello ECOC SVM. Le features utilizzate per l'addestramento sono state estratte dal penultimo strato convoluzionale fully connected (*fc7*) di AlexNet.

## Risultati

### Fase 1

Il primo addestramento della rete neurale è risultato in un evidente underfitting: l'addestramento non ha portato a nessun miglioramento dell'accuratezza, come mostrato dal grafico in figura (1). Queste prestazioni sono paragonabili a quelle di un classificatore casuale.

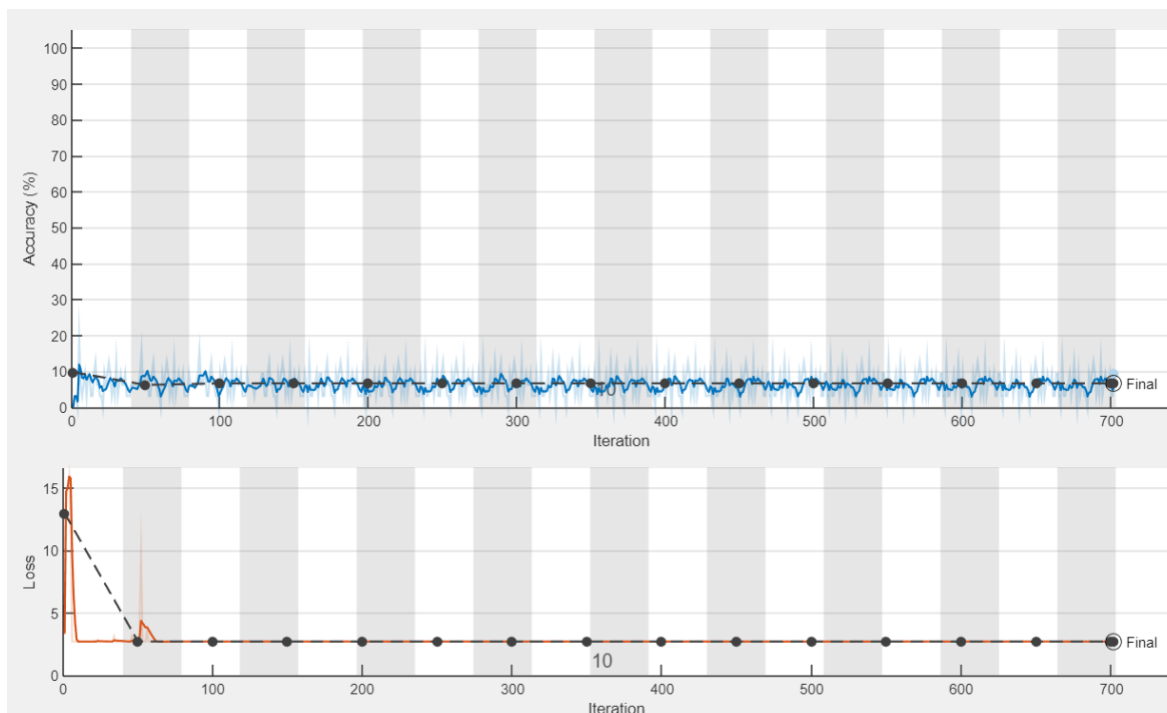


Figura 1



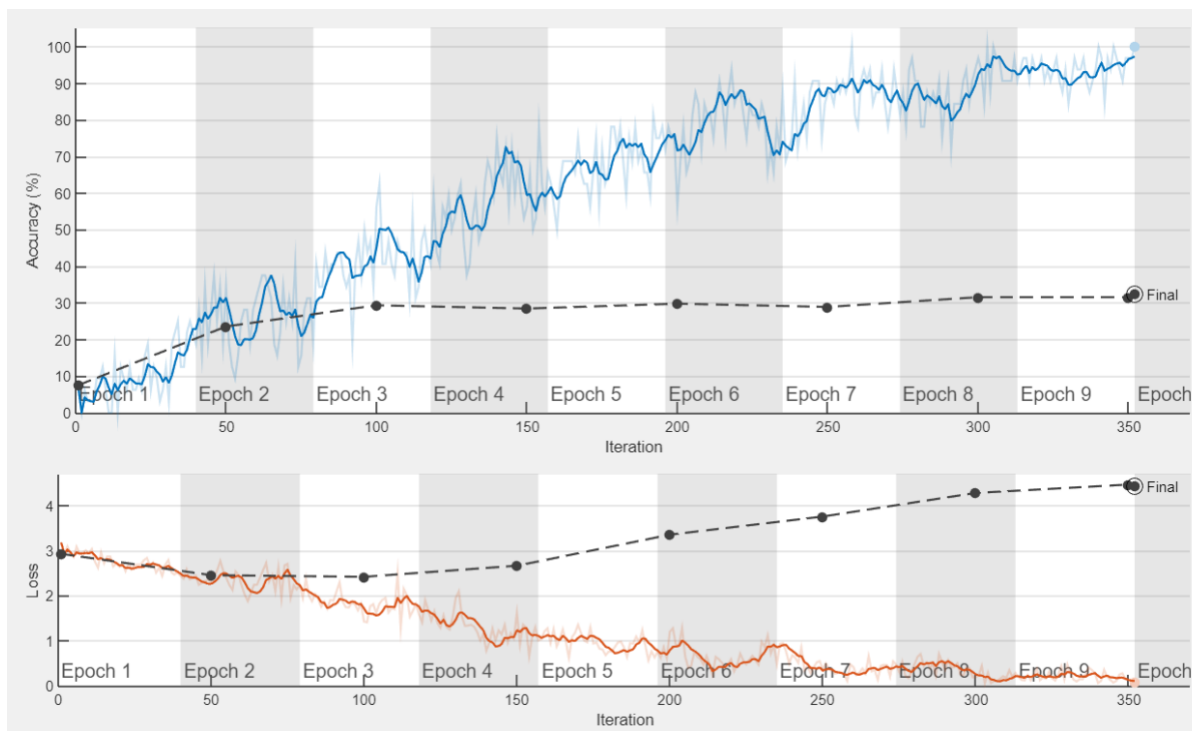


Figura 3

		Confusion Matrix																
		Bedroom	Coast	Forest	Highway	Industrial	InsideCity	Kitchen	LivingRoom	Mountain	Office	OpenCountry	Store	Street	Suburb	TallBuilding		
Output Class	Bedroom	17	4	12	2	12	3	13	27	14	10	0	7	4	4	27	10.9%	
		0.6%	0.1%	0.4%	0.1%	0.4%	0.1%	0.4%	0.9%	0.5%	0.3%	0.0%	0.2%	0.1%	0.1%	0.9%	89.1%	
	Coast	10	113	4	25	6	3	1	3	39	4	65	1	4	3	1	40.1%	
		0.3%	3.8%	0.1%	0.8%	0.2%	0.1%	0.0%	0.1%	1.3%	0.1%	2.2%	0.0%	0.1%	0.1%	0.0%	59.9%	
	Forest	0	2	30	1	3	9	3	2	6	2	4	9	2	3	4	37.5%	
		0.0%	0.1%	1.0%	0.0%	0.1%	0.3%	0.1%	0.1%	0.2%	0.1%	0.1%	0.3%	0.1%	0.1%	0.1%	62.5%	
	Highway	1	37	1	97	23	7	5	5	10	2	30	0	4	1	2	43.1%	
		0.0%	1.2%	0.0%	3.2%	0.8%	0.2%	0.2%	0.2%	0.3%	0.1%	1.0%	0.0%	0.1%	0.0%	0.1%	56.9%	
	Industrial	8	3	14	3	40	13	11	8	15	6	5	11	8	9	22	22.7%	
		0.3%	0.1%	0.5%	0.1%	1.3%	0.4%	0.4%	0.3%	0.5%	0.2%	0.2%	0.4%	0.3%	0.3%	0.7%	77.3%	
	InsideCity	4	4	17	2	10	40	3	4	10	4	7	20	11	8	7	26.5%	
		0.1%	0.1%	0.6%	0.1%	0.3%	1.3%	0.1%	0.1%	0.3%	0.1%	0.2%	0.7%	0.4%	0.3%	0.2%	73.5%	
	Kitchen	19	1	5	0	14	20	26	42	5	9	4	25	7	6	22	12.7%	
		0.6%	0.0%	0.2%	0.0%	0.5%	0.7%	0.9%	1.4%	0.2%	0.3%	0.1%	0.8%	0.2%	0.2%	0.7%	87.3%	
	LivingRoom	12	0	6	1	4	3	8	35	6	8	3	10	4	6	18	26.2%	
		0.4%	0.0%	0.2%	0.0%	0.1%	0.1%	0.3%	1.2%	0.2%	0.3%	0.1%	0.3%	0.1%	0.2%	0.6%	71.8%	
	Mountain	18	42	54	8	27	22	7	18	109	11	91	18	16	14	19	23.0%	
	0.6%	1.4%	1.8%	0.3%	0.9%	0.7%	0.2%	0.6%	3.7%	0.4%	3.0%	0.6%	0.5%	0.5%	0.6%	77.0%		
Office	5	7	25	2	19	13	5	9	12	47	8	16	6	2	41	21.7%		
	0.2%	0.2%	0.8%	0.1%	0.6%	0.4%	0.2%	0.3%	0.4%	1.6%	0.3%	0.5%	0.2%	0.1%	1.4%	78.3%		
OpenCountry	0	42	12	8	10	15	4	4	27	5	68	6	7	5	2	31.6%		
	0.0%	1.4%	0.4%	0.3%	0.3%	0.5%	0.1%	0.1%	0.9%	0.2%	2.3%	0.2%	0.2%	0.2%	0.1%	68.4%		
Store	3	2	23	4	12	33	7	7	4	3	10	70	6	5	12	34.8%		
	0.1%	0.1%	0.8%	0.1%	0.4%	1.1%	0.2%	0.2%	0.1%	0.1%	0.3%	2.3%	0.2%	0.2%	0.4%	65.2%		
Street	10	1	12	6	11	12	7	12	9	0	10	13	101	19	9	43.5%		
	0.3%	0.0%	0.4%	0.2%	0.4%	0.4%	0.2%	0.4%	0.3%	0.0%	0.3%	0.4%	3.4%	0.6%	0.3%	56.5%		
Suburb	3	1	5	1	8	8	5	4	7	1	5	4	12	55	5	44.4%		
	0.1%	0.0%	0.2%	0.0%	0.3%	0.3%	0.2%	0.1%	0.2%	0.0%	0.2%	0.1%	0.4%	1.8%	0.2%	55.6%		
TallBuilding	6	1	8	0	12	7	5	9	1	3	0	5	0	1	65	52.8%		
	0.2%	0.0%	0.3%	0.0%	0.4%	0.2%	0.2%	0.3%	0.0%	0.1%	0.0%	0.2%	0.0%	0.0%	2.2%	47.2%		
		14.7%	43.5%	13.2%	60.6%	19.0%	19.2%	23.6%	18.5%	39.8%	40.9%	21.9%	32.6%	52.6%	39.0%	25.4%	30.6%	
		85.3%	56.5%	86.8%	39.4%	81.0%	80.8%	76.4%	81.5%	60.2%	59.1%	78.1%	67.4%	47.4%	61.0%	74.6%	69.4%	
		Target Class																

Figura 4

Diminuendo ulteriormente il parametro *InitialLearnRate* fino a 0.0005, è diminuita l'accuratezza sull'insieme di test (27% circa), mentre quella misurata sull'insieme di validazione è rimasta sostanzialmente invariata (32%circa). Questo suggerisce un leggero overfitting dei dati che è stato opportunamente affrontato nella fase 2.

## Fase 2

Di seguito sono descritti i vari tentativi di miglioramento della baseline.

### Data augmentation

Il primo tentativo di miglioramento con l'impiego della data augmentation (left-to-right reflection) è stato soddisfacente e ha portato al raggiungimento di un'accuratezza sui dati di test di circa il 40% (+10%).

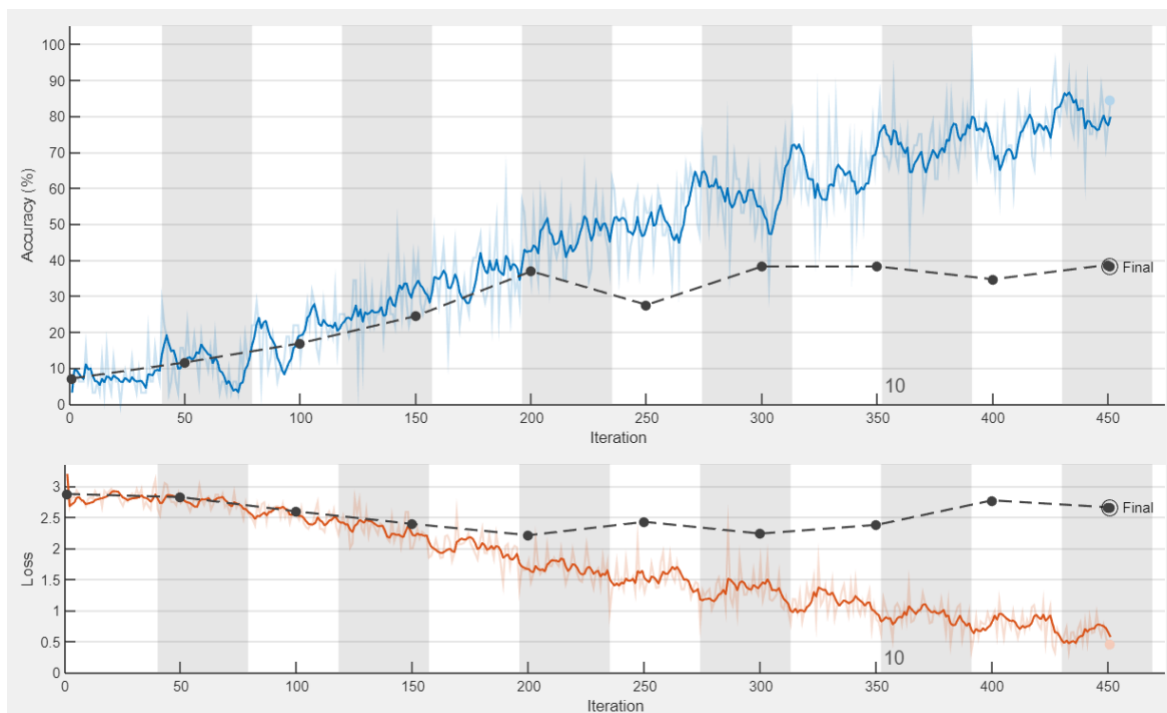


Figura 5



		Confusion Matrix																
Output Class	Bedroom	40	2	6	7	17	11	30	55	21	15	8	6	10	8	18	15.7%	
		1.3%	0.1%	0.2%	0.2%	0.6%	0.4%	1.0%	1.8%	0.7%	0.5%	0.3%	0.2%	0.3%	0.3%	0.6%	84.3%	
	Coast	3	180	9	19	6	5	3	1	52	1	108	3	1	3	0	45.7%	
		0.1%	6.0%	0.3%	0.6%	0.2%	0.2%	0.1%	0.0%	1.7%	0.0%	3.6%	0.1%	0.0%	0.1%	0.0%	54.3%	
	Forest	1	2	103	1	9	14	1	1	8	0	3	21	5	4	2	58.9%	
		0.0%	0.1%	3.5%	0.0%	0.3%	0.5%	0.0%	0.0%	0.3%	0.0%	0.1%	0.7%	0.2%	0.1%	0.1%	41.1%	
	Highway	2	33	5	100	8	2	2	3	18	3	39	2	8	4	0	43.7%	
		0.1%	1.1%	0.2%	3.4%	0.3%	0.1%	0.1%	0.1%	0.6%	0.1%	1.3%	0.1%	0.3%	0.1%	0.0%	56.3%	
	Industrial	5	0	2	2	36	7	8	8	7	3	3	6	5	1	11	34.6%	
		0.2%	0.0%	0.1%	0.1%	1.2%	0.2%	0.3%	0.3%	0.2%	0.1%	0.1%	0.2%	0.2%	0.0%	0.4%	65.4%	
	InsideCity	2	1	7	5	19	46	4	5	12	1	6	20	13	4	12	29.3%	
		0.1%	0.0%	0.2%	0.2%	0.6%	1.5%	0.1%	0.2%	0.4%	0.0%	0.2%	0.7%	0.4%	0.1%	0.4%	70.7%	
	Kitchen	8	0	1	0	7	5	23	19	7	0	2	5	3	0	2	28.0%	
		0.3%	0.0%	0.0%	0.0%	0.2%	0.2%	0.8%	0.6%	0.2%	0.0%	0.1%	0.2%	0.1%	0.0%	0.1%	72.0%	
	LivingRoom	22	0	1	0	8	7	6	48	3	2	5	6	2	5	10	38.4%	
		0.7%	0.0%	0.0%	0.0%	0.3%	0.2%	0.2%	1.6%	0.1%	0.1%	0.2%	0.2%	0.1%	0.2%	0.3%	61.6%	
	Mountain	4	6	16	1	11	10	5	3	71	2	34	11	5	5	6	37.4%	
		0.1%	0.2%	0.5%	0.0%	0.4%	0.3%	0.2%	0.1%	2.4%	0.1%	1.1%	0.4%	0.2%	0.2%	0.2%	62.6%	
	Office	6	6	28	4	15	11	9	11	27	74	4	14	5	1	34	29.7%	
		0.2%	0.2%	0.9%	0.1%	0.5%	0.4%	0.3%	0.4%	0.9%	2.5%	0.1%	0.5%	0.2%	0.0%	1.1%	70.3%	
OpenCountry	0	27	6	12	11	13	1	1	25	1	84	4	6	7	2	42.0%		
	0.0%	0.9%	0.2%	0.4%	0.4%	0.4%	0.0%	0.0%	0.8%	0.0%	2.8%	0.1%	0.2%	0.2%	0.1%	58.0%		
Store	2	0	34	1	10	35	8	11	8	2	4	94	7	8	8	40.5%		
	0.1%	0.0%	1.1%	0.0%	0.3%	1.2%	0.3%	0.4%	0.3%	0.1%	0.1%	3.1%	0.2%	0.3%	0.3%	59.5%		
Street	6	0	1	5	9	9	2	9	1	2	1	2	106	8	0	65.8%		
	0.2%	0.0%	0.0%	0.2%	0.3%	0.3%	0.1%	0.3%	0.0%	0.1%	0.0%	0.1%	3.6%	0.3%	0.0%	34.2%		
Suburb	7	3	5	3	11	18	5	4	7	4	9	7	13	80	2	44.9%		
	0.2%	0.1%	0.2%	0.1%	0.4%	0.6%	0.2%	0.1%	0.2%	0.1%	0.3%	0.2%	0.4%	2.7%	0.1%	55.1%		
TallBuilding	8	0	4	0	34	15	3	10	7	5	0	14	3	3	149	58.4%		
	0.3%	0.0%	0.1%	0.0%	1.1%	0.5%	0.1%	0.3%	0.2%	0.2%	0.0%	0.5%	0.1%	0.1%	5.0%	41.6%		
		34.5%	69.2%	45.2%	62.5%	17.1%	22.1%	20.9%	25.4%	25.9%	64.3%	27.1%	43.7%	55.2%	56.7%	58.2%	41.3%	
		65.5%	30.8%	54.8%	37.5%	82.9%	77.9%	79.1%	74.6%	74.1%	35.7%	72.9%	56.3%	44.8%	43.3%	41.8%	58.7%	
		Bedroom	Coast	Forest	Highway	Industrial	InsideCity	Kitchen	LivingRoom	Mountain	Office	OpenCountry	Store	Street	Suburb	TallBuilding		
		Target Class																

Figura 6

## Batch normalization

L'introduzione della batch normalization [Ioffe and Szegedy, 2015] ha portato a ulteriori significativi miglioramenti, sotto diversi punti di vista. Innanzitutto, l'accuratezza sui dati di test è aumentata fino al 48% circa (+8%), mentre la perdita misurata sui dati di validazione è scesa fino a 1.7 circa. Questo significa che non solo ci sono più classificazioni corrette, ma anche che c'è una maggior confidenza nelle stesse. Inoltre, dal grafico in figura (7), si nota subito una durata di addestramento maggiore rispetto ai casi precedenti. Questo è dovuto al fatto che la perdita sull'insieme di validazione ha un andamento decrescente e di conseguenza il criterio di early stopping basato sul parametro *ValidationPatience* non è mai stato soddisfatto. L'addestramento si è comunque concluso automaticamente al raggiungimento della trentesima epoca essendo 30 il valore di default del parametro *MaxEpochs*.

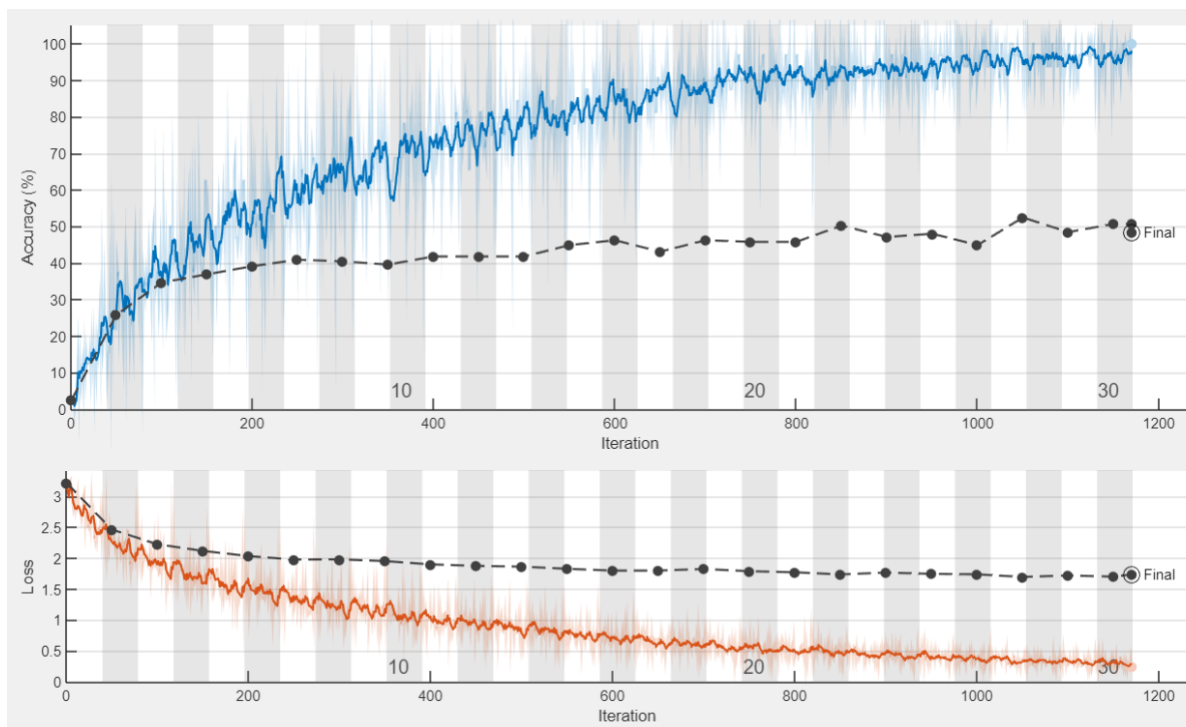


Figura 7

		Confusion Matrix																
Output Class	Bedroom	47	3	4	1	14	6	28	38	12	13	4	4	3	3	17	23.9%	
		1.6%	0.1%	0.1%	0.0%	0.5%	0.2%	0.9%	1.3%	0.4%	0.4%	0.1%	0.1%	0.1%	0.1%	0.6%	76.1%	
	Coast	1	172	8	22	11	1	3	2	42	1	103	0	0	2	2	46.5%	
		0.0%	5.8%	0.3%	0.7%	0.4%	0.0%	0.1%	0.1%	1.4%	0.0%	3.5%	0.0%	0.0%	0.1%	0.1%	53.5%	
	Forest	1	3	127	2	9	21	1	0	11	0	8	43	13	5	11	49.8%	
		0.0%	0.1%	4.3%	0.1%	0.3%	0.7%	0.0%	0.0%	0.4%	0.0%	0.3%	1.4%	0.4%	0.2%	0.4%	50.2%	
	Highway	2	31	2	108	4	6	2	0	8	0	25	0	6	0	1	55.4%	
		0.1%	1.0%	0.1%	3.6%	0.1%	0.2%	0.1%	0.0%	0.3%	0.0%	0.8%	0.0%	0.2%	0.0%	0.0%	44.6%	
	Industrial	14	0	11	4	76	9	8	18	20	10	6	7	15	6	23	33.5%	
		0.5%	0.0%	0.4%	0.1%	2.5%	0.3%	0.3%	0.6%	0.7%	0.3%	0.2%	0.2%	0.5%	0.2%	0.8%	66.5%	
	InsideCity	1	2	1	1	12	53	5	7	7	6	5	5	5	9	5	42.7%	
		0.0%	0.1%	0.0%	0.0%	0.4%	1.8%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.3%	0.2%	57.3%	
	Kitchen	18	1	2	0	7	16	31	29	2	5	3	7	8	5	10	21.5%	
		0.6%	0.0%	0.1%	0.0%	0.2%	0.5%	1.0%	1.0%	0.1%	0.2%	0.1%	0.2%	0.3%	0.2%	0.3%	78.5%	
	LivingRoom	11	0	0	0	4	3	8	50	0	0	1	6	1	3	13	50.0%	
		0.4%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	1.7%	0.0%	0.0%	0.0%	0.2%	0.0%	0.1%	0.4%	50.0%	
	Mountain	5	15	32	6	12	8	2	2	115	1	24	5	4	1	2	49.1%	
		0.2%	0.5%	1.1%	0.2%	0.4%	0.3%	0.1%	0.1%	3.9%	0.0%	0.8%	0.2%	0.1%	0.0%	0.1%	50.9%	
	Office	1	0	3	0	5	7	6	5	4	68	0	1	0	2	5	53.6%	
		0.0%	0.0%	0.1%	0.0%	0.2%	0.2%	0.2%	0.2%	0.1%	2.3%	0.0%	0.0%	0.0%	0.1%	0.2%	36.4%	
	OpenCountry	0	33	10	10	8	13	1	1	31	1	118	0	5	0	5	50.0%	
		0.0%	1.1%	0.3%	0.3%	0.3%	0.4%	0.0%	0.0%	1.0%	0.0%	4.0%	0.0%	0.2%	0.0%	0.2%	50.0%	
	Store	6	0	9	1	17	38	11	18	5	1	2	117	6	14	5	46.8%	
		0.2%	0.0%	0.3%	0.0%	0.6%	1.3%	0.4%	0.6%	0.2%	0.0%	0.1%	3.9%	0.2%	0.5%	0.2%	53.2%	
	Street	3	0	11	5	6	11	1	7	3	1	6	4	118	7	2	53.8%	
		0.1%	0.0%	0.4%	0.2%	0.2%	0.4%	0.0%	0.2%	0.1%	0.0%	0.2%	0.1%	4.0%	0.2%	0.1%	36.2%	
	Suburb	1	0	2	0	4	9	0	3	6	2	4	9	5	83	0	54.8%	
		0.0%	0.0%	0.1%	0.0%	0.1%	0.3%	0.0%	0.1%	0.2%	0.1%	0.1%	0.3%	0.2%	2.8%	0.0%	35.2%	
	TallBuilding	5	0	6	0	22	7	3	9	8	6	1	7	3	1	155	36.5%	
		0.2%	0.0%	0.2%	0.0%	0.7%	0.2%	0.1%	0.3%	0.3%	0.2%	0.0%	0.2%	0.1%	0.0%	5.2%	33.5%	
		40.5%	66.2%	55.7%	67.5%	36.0%	25.5%	28.2%	26.5%	42.0%	59.1%	38.1%	54.4%	61.5%	58.9%	30.5%	48.2%	
		59.5%	33.8%	44.3%	32.5%	64.0%	74.5%	71.8%	73.5%	58.0%	40.9%	61.9%	45.6%	38.5%	41.1%	39.5%	51.8%	
		Bedroom	Coast	Forest	Highway	Industrial	InsideCity	Kitchen	LivingRoom	Mountain	Office	OpenCountry	Store	Street	Suburb	TallBuilding		
		Target Class																

Figura 8

Grazie all'introduzione della batch normalization è stato inoltre possibile incrementare il valore di *InitialLearnRate* fino a 0.001 mantenendo le prestazioni pressoché inalterate (accuratezza sui dati di test pari al 48% circa) ma diminuendo il tempo di addestramento.

## Incremento del numero e della dimensione dei filtri

Mantenendo inalterato il tasso di apprendimento (pari a 0.001), l'incremento del numero e della dimensione dei filtri ha portato l'accuratezza sui dati di test a raggiungere il 54% circa (+6%).

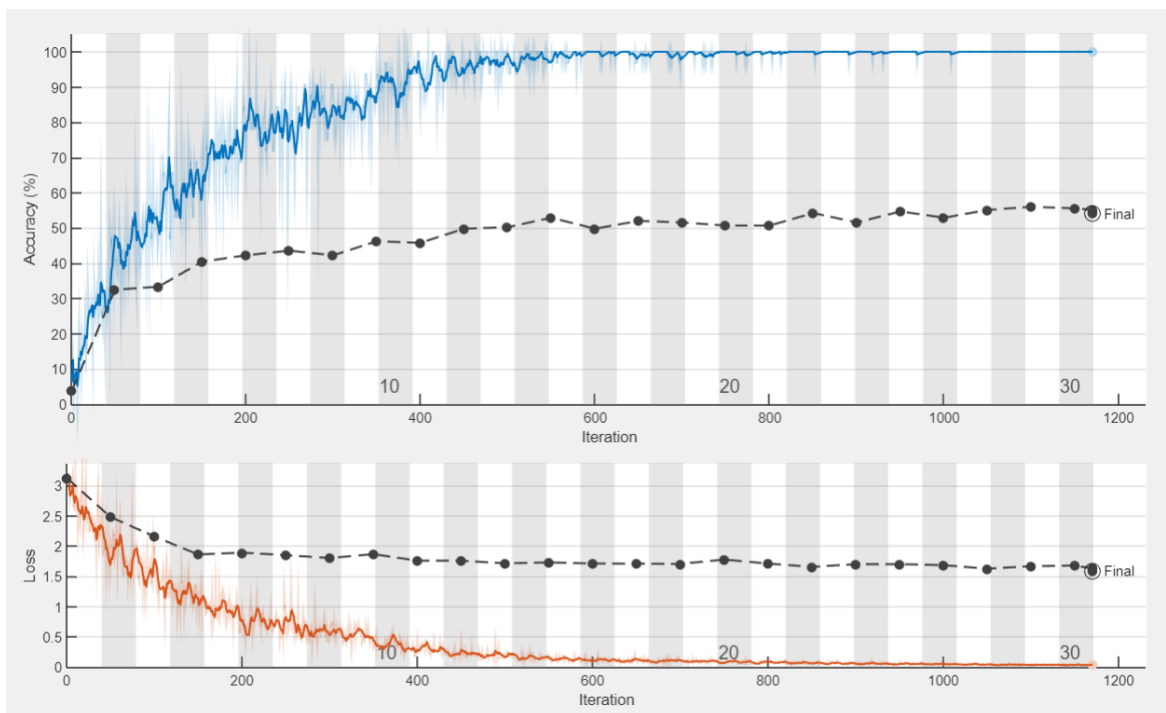


Figura 9

		Confusion Matrix																Output Class
		Bedroom	Coast	Forest	Highway	Industrial	InsideCity	Kitchen	LivingRoom	Mountain	Office	OpenCountry	Store	Street	Suburb	TallBuilding		
Output Class	Bedroom	52	0	3	0	14	3	22	31	10	9	1	7	7	3	9	30.4%	
		1.7%	0.0%	0.1%	0.0%	0.5%	0.1%	0.7%	1.0%	0.3%	0.3%	0.0%	0.2%	0.2%	0.1%	0.3%	69.6%	
	Coast	2	159	1	12	2	1	1	1	21	0	71	0	0	0	0	58.7%	
		0.1%	5.3%	0.0%	0.4%	0.1%	0.0%	0.0%	0.0%	0.7%	0.0%	2.4%	0.0%	0.0%	0.0%	0.0%	41.3%	
	Forest	1	2	168	4	10	16	2	2	19	3	9	34	15	1	6	57.5%	
		0.0%	0.1%	5.6%	0.1%	0.3%	0.5%	0.1%	0.1%	0.6%	0.1%	0.3%	1.1%	0.5%	0.0%	0.2%	42.5%	
	Highway	2	43	0	114	8	7	3	1	11	0	24	0	6	1	0	51.8%	
		0.1%	1.4%	0.0%	3.8%	0.3%	0.2%	0.1%	0.0%	0.4%	0.0%	0.8%	0.0%	0.2%	0.0%	0.0%	48.2%	
	Industrial	9	0	10	1	76	10	6	17	12	11	0	7	10	7	16	39.6%	
		0.3%	0.0%	0.3%	0.0%	2.5%	0.3%	0.2%	0.6%	0.4%	0.4%	0.0%	0.2%	0.3%	0.2%	0.5%	60.4%	
	InsideCity	4	1	1	2	13	75	5	4	5	6	9	18	5	7	9	45.7%	
		0.1%	0.0%	0.0%	0.1%	0.4%	2.5%	0.2%	0.1%	0.2%	0.2%	0.3%	0.6%	0.2%	0.2%	0.3%	54.3%	
	Kitchen	16	0	1	1	5	10	34	34	1	3	2	5	1	2	5	28.3%	
		0.5%	0.0%	0.0%	0.0%	0.3%	1.1%	1.1%	1.1%	0.0%	0.1%	0.1%	0.2%	0.0%	0.1%	0.2%	71.7%	
	LivingRoom	15	0	0	0	3	1	11	62	0	5	3	2	2	2	14	51.7%	
		0.5%	0.0%	0.0%	0.0%	0.1%	0.0%	0.4%	2.1%	0.0%	0.2%	0.1%	0.1%	0.1%	0.1%	0.5%	48.3%	
	Mountain	3	6	13	6	6	2	1	3	129	0	18	3	3	2	1	55.8%	
		0.1%	0.2%	0.4%	0.2%	0.2%	0.1%	0.0%	0.1%	4.3%	0.0%	0.6%	0.1%	0.1%	0.1%	0.0%	44.2%	
	Office	3	0	2	1	12	7	10	7	1	62	1	4	1	0	12	50.4%	
		0.1%	0.0%	0.1%	0.0%	0.4%	0.2%	0.3%	0.2%	0.0%	2.1%	0.0%	0.1%	0.0%	0.0%	0.4%	49.6%	
	OpenCountry	0	47	4	16	10	10	1	3	46	0	164	2	6	2	2	52.4%	
		0.0%	1.6%	0.1%	0.5%	0.3%	0.3%	0.0%	0.1%	1.5%	0.0%	5.5%	0.1%	0.2%	0.1%	0.1%	47.6%	
	Store	0	0	15	0	16	30	6	7	4	0	0	113	4	4	2	56.2%	
		0.0%	0.0%	0.5%	0.0%	0.5%	1.0%	0.2%	0.2%	0.1%	0.0%	0.0%	3.8%	0.1%	0.1%	0.1%	43.8%	
	Street	1	0	7	3	4	10	0	2	3	0	4	3	122	2	1	75.3%	
		0.0%	0.0%	0.2%	0.1%	0.1%	0.3%	0.0%	0.1%	0.1%	0.0%	0.1%	0.1%	4.1%	0.1%	0.0%	24.7%	
	Suburb	1	1	1	0	7	11	4	3	7	3	3	7	9	107	0	65.2%	
		0.0%	0.0%	0.0%	0.0%	0.2%	0.4%	0.1%	0.1%	0.2%	0.1%	0.1%	0.2%	0.3%	3.6%	0.0%	34.8%	
	TallBuilding	7	1	2	0	25	15	4	12	5	13	1	10	1	1	179	54.9%	
		0.2%	0.0%	0.1%	0.0%	0.8%	0.5%	0.1%	0.4%	0.2%	0.4%	0.0%	0.3%	0.0%	0.0%	6.0%	45.1%	
		44.8%	61.2%	73.7%	71.2%	36.0%	36.1%	30.9%	32.8%	47.1%	53.9%	52.9%	52.6%	63.5%	75.9%	69.9%	54.1%	
		55.2%	38.8%	26.3%	28.7%	64.0%	63.9%	69.1%	67.2%	52.9%	46.1%	47.1%	47.4%	36.5%	24.1%	30.1%	45.9%	
		Target Class																
		Bedroom	Coast	Forest	Highway	Industrial	InsideCity	Kitchen	LivingRoom	Mountain	Office	OpenCountry	Store	Street	Suburb	TallBuilding		



Figura 10

### Incremento della dimensione del minibatch

La dimensione del minibatch è stata aumentata inizialmente a 64 e successivamente a 128. In entrambi i casi le prestazioni sono rimaste sostanzialmente invariate. Ciononostante è stato deciso di mantenere il valore 128 in modo da avere una miglior stima del gradiente anche durante i successivi tentativi di miglioramento.

### Incremento del numero di strati convoluzionali e dropout

L'obiettivo della fase 2 è stato infine raggiunto introducendo due strati convoluzionali, uno strato di dropout e modificando alcune opzioni di apprendimento. In particolare, il primo strato convoluzionale introdotto è costituito da 64 filtri 9x9 mentre il secondo da 64 filtri 11x11. Lo strato di dropout è stato inserito prima dell'ultimo strato completamente connesso. L'incremento di *InitialLearnRate* a 0.005 e di *MaxEpochs* a 50 ha permesso il raggiungimento delle migliori prestazioni possibili con questa nuova configurazione: l'accuratezza sui dati di test ha raggiunto il 62% circa (+4%) come mostrato nel grafico in figura 12.

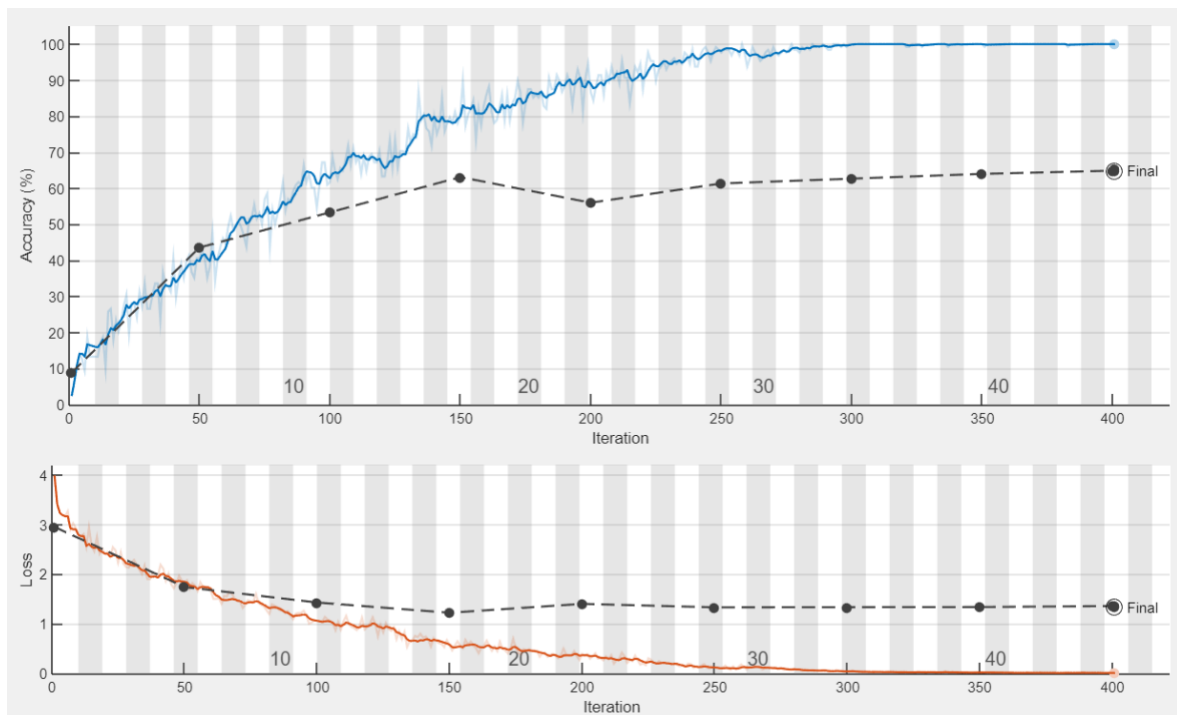


Figura 11

		Confusion Matrix																
Output Class	Bedroom	60 2.0%	0 0.0%	2 0.1%	1 0.0%	7 0.2%	2 0.1%	14 0.5%	42 1.4%	10 0.3%	7 0.2%	1 0.0%	3 0.1%	1 0.0%	2 0.1%	8 0.3%	37.5%	
	Coast	2 0.1%	164 5.5%	0 0.0%	21 0.7%	3 0.1%	0 0.0%	0 0.0%	0 0.0%	11 0.4%	1 0.0%	49 1.6%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	64.8%	
	Forest	2 0.1%	1 0.0%	174 5.8%	4 0.1%	9 0.3%	10 0.3%	1 0.0%	1 0.0%	14 0.5%	2 0.1%	10 0.3%	9 0.3%	6 0.2%	0 0.0%	6 0.2%	69.9%	
	Highway	2 0.1%	25 0.8%	0 0.0%	112 3.8%	5 0.2%	3 0.1%	2 0.1%	1 0.0%	6 0.2%	0 0.0%	14 0.5%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	65.1%	
	Industrial	4 0.1%	0 0.0%	7 0.2%	3 0.1%	85 2.8%	11 0.4%	7 0.2%	16 0.5%	9 0.3%	4 0.1%	4 0.1%	7 0.2%	11 0.4%	4 0.1%	21 0.7%	44.0%	
	InsideCity	2 0.1%	0 0.0%	0 0.0%	1 0.0%	14 0.5%	99 3.3%	1 0.0%	2 0.1%	1 0.0%	2 0.1%	5 0.2%	14 0.5%	8 0.3%	5 0.2%	5 0.2%	62.3%	
	Kitchen	12 0.4%	0 0.0%	0 0.0%	3 0.1%	9 0.3%	18 0.6%	57 1.9%	28 0.9%	0 0.0%	3 0.1%	3 0.1%	8 0.3%	1 0.1%	3 0.1%	4 0.1%	38.3%	
	LivingRoom	19 0.6%	0 0.0%	2 0.1%	1 0.0%	8 0.3%	1 0.0%	5 0.2%	72 2.4%	0 0.0%	10 0.3%	2 0.1%	1 0.0%	3 0.1%	4 0.1%	5 0.2%	54.1%	
	Mountain	3 0.1%	10 0.3%	20 0.7%	2 0.1%	12 0.4%	1 0.0%	1 0.0%	0 0.0%	189 6.3%	1 0.0%	28 0.9%	5 0.2%	2 0.1%	1 0.0%	3 0.1%	68.0%	
	Office	6 0.2%	1 0.0%	0 0.0%	0 0.0%	8 0.3%	5 0.2%	7 0.2%	7 0.2%	2 0.1%	80 2.7%	2 0.1%	2 0.1%	0 0.0%	1 0.0%	9 0.3%	61.5%	
	OpenCountry	0 0.0%	57 1.9%	3 0.1%	8 0.3%	10 0.3%	9 0.3%	0 0.0%	0 0.0%	21 0.7%	1 0.0%	188 6.3%	2 0.1%	1 0.0%	2 0.1%	1 0.0%	62.0%	
	Store	1 0.0%	0 0.0%	13 0.4%	2 0.1%	13 0.4%	27 0.9%	5 0.2%	8 0.3%	2 0.1%	0 0.0%	1 0.0%	142 4.8%	12 0.4%	5 0.2%	2 0.1%	60.9%	
	Street	1 0.0%	0 0.0%	4 0.1%	1 0.0%	3 0.1%	4 0.1%	2 0.1%	1 0.0%	2 0.1%	0 0.0%	1 0.0%	5 0.2%	136 4.6%	0 0.0%	3 0.1%	83.4%	
	Suburb	0 0.0%	2 0.1%	0 0.0%	1 0.0%	4 0.1%	8 0.3%	4 0.1%	3 0.1%	3 0.1%	1 0.0%	2 0.1%	5 0.2%	6 0.2%	112 3.8%	0 0.0%	73.7%	
	TallBuilding	2 0.1%	0 0.0%	2 0.1%	0 0.0%	21 0.7%	10 0.3%	4 0.1%	7 0.2%	4 0.1%	3 0.1%	0 0.0%	11 0.4%	3 0.1%	2 0.1%	189 6.3%	73.3%	
			51.7%	63.1%	76.3%	70.0%	40.3%	47.6%	51.8%	38.1%	69.0%	69.6%	60.6%	66.0%	70.8%	79.4%	73.8%	62.3%
			48.3%	36.9%	23.7%	30.0%	59.7%	52.4%	48.2%	61.9%	31.0%	30.4%	39.4%	34.0%	29.2%	20.6%	26.2%	37.7%
		Bedroom	Coast	Forest	Highway	Industrial	InsideCity	Kitchen	LivingRoom	Mountain	Office	OpenCountry	Store	Street	Suburb	TallBuilding		
		Target Class																

Figura 12

## Fase 3

Come atteso, entrambi gli approcci di transfer learning proposti hanno portato a importanti miglioramenti delle prestazioni del classificatore. L'accuratezza raggiunta sui dati di test è stata, in entrambi i casi, dell'86% circa. Per quanto riguarda l'approccio "feature extraction" sono state rilevate prestazioni pressoché identiche sia utilizzando un classificatore PAIRWISE SVM che ECOC SVM (accuratezza sui dati di test pari a 86% circa).

# Riferimenti

- Codice: <https://github.com/heapify00/cv-project>
- [Ioffe and Szegedy, 2015] Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*